Impacts of Climate Change on Discharge in Switzerland: Cascading Uncertainties and Robustness in Models

Nans Addor

PhD thesis - University of Zurich

Cover picture: The Verzasca, descending the southern flanks of the Swiss Alps, is one of the six rivers whose discharge under climate change has been investigated in this thesis. Projections indicate with confidence a significant decrease of summer discharge, an impact that could be halved if considerable efforts to reduce greenhouse gas emissions were undertaken. Photograph by Nans Addor.

Impacts of Climate Change on Discharge in Switzerland: Cascading Uncertainties and Robustness in Models

Dissertation

zur

Erlangung der naturwissenschaftlichen Doktorwürde

(Dr. sc. nat.)

vorgelegt der

Mathematisch-naturwissenschaftlichen Fakultät

der

Universität Zürich

von

Nans Antoine Philippe Addor

von

Sainte-Croix VD

Promotionskomitee

Prof. Dr. Jan Seibert (Vorsitz)

Prof. Dr. Reinhard Furrer

Prof. Dr. Reto Knutti

Zürich, 2015

'[...] models' usefulness rests in their ability to help us think about a system - as heuristic tools to help us understand what we can actually observe or estimate, or as a way to challenge existing formulations and intuitions.'

Weaver et al. (2013)

Summary

Climate change is ongoing, but although the occurrence of temperature increase at the planetary scale is well established, its impacts at the watershed scale are still unclear. Impacts on river discharge are of particular relevance for society, because of the importance of rivers, for instance for water resources, ecosystems, energy production and recreation. The investigation of hydrological changes induced by global warming relies principally on numerical models, which simulate climate evolution and its consequences on watershed hydrology. Yet since models are, by nature, simplified representations of the reality, it is crucial to account for their abilities and limitations when interpreting their simulations. This implies that the uncertainties in climate and discharge projections need to be characterised. These uncertainties result from different sources, including future greenhouse gas emissions, the chaotic nature of the atmosphere, our understanding of the system and its representation in models, errors in observational data sets and the vulnerability of society to climatic and hydrological events. Although these uncertainties are not necessarily reducible or quantifiable, their exploration is key in order to improve our understanding of climate change and its impacts, inform future model development and support decision-making on adaptation strategies. In parallel to the characterisation of uncertainties related to climate change, there is an ongoing quest for more robustness: for robust models, reliable even under changing climatic conditions; for robust ensemble projections, that agree on significant changes despite their uncertainty; for robust adaptation measures, that benefit society for a wide range of climate outcomes.

This thesis shows that despite uncertainties in discharge projections in Switzerland, robust changes emerge, such as a decrease of summer discharge and earlier snow and icemelt peak. These results emerge from on a particularly large ensemble of model chains, which also reveals that the projected impacts on hydrological regimes could be reduced by about a factor two, if stringent efforts to reduce emissions were undertaken. When combining climate model and hydrological models, dealing with systematic differences between climate model simulations of precipitation and temperature and their observations poses a considerable challenge. These differences, usually referred to as 'biases', can induce anomalies in hydrological simulations. It is shown here that part of these biases at the regional scale (over Switzerland) stem from the representation of the atmospheric circulation at the synoptic scale (over Europe), which compromises our capacity to reduces these biases by post-processing. Further, it is demonstrated that these biases do not stem from climate model limitations alone, but also from uncertainties related to observational data sets and natural variability of the climate system. These results underscore the importance of identifying and accounting for the causes of biases in order to progress towards more robust bias-correction techniques. In the context of hydrological modeling, a novel method to assess the robustness of hydrological models under climate change is proposed. It relies on how well observed trends in discharge and evapotranspiration over the last decades are simulated. It can be used to assess the potential benefits of time-dependent model parameter values, over a standard setup relying on constant parameter values. Finally, it is proposed that better projections alone are not sufficient to significantly progress with the design of adaptation strategies. Additional steps are necessary, in particular the fostering of interdisciplinary dialogue between scientists and decision-makers, and, in academia, the exposure of students to uncertainties inherent to climate change, and ways to deal with them. This idea was put into practice by organizing a multi-day workshop at the University of Zurich. Its main outcomes and recommendations for future similar events are presented. In conclusion, exploring abilities and limitations of models, and engaging in interdisciplinary dialogue, is essential to benefit from the full potential of models in the context of climate change.

Zusammenfassung

Das Klima ändert sich, aber obwohl man sich darüber einig ist, dass die Temperatur global ansteigt, sind die Folgen des Klimawandels auf regionaler Ebene vielfach noch unklar. Für die Gesellschaft sind die Auswirkungen des Klimawandels auf den Wasserkreislauf von besonderer Bedeutung, denn Wasserressourcen, Stromproduktion, Ökosysteme und Freizeitaktivitäten etc. stehen im direkten Zusammenhang. Um hydrologische Veränderungen bedingt durch den Klimawandel auf der Einzugsgebietsebene zu untersuchen, werden hauptsächlich numerische Modelle verwendet. Weil Modelle jedoch eine vereinfachte Repräsentation der Realität sind, ist es wichtig deren Möglichkeiten und Grenzen bei Analyse und Interpretation der Modellresultate zu berücksichtigen. Dafür müssen Unsicherheiten in Klima- und Abflusssimulationen charakterisiert werden. Unsicherheiten haben unterschiedliche Ursachen, unter anderem zukünftige Treibhausgasemissionen, das chaotische Verhalten der Atmosphäre, unser Verständnis des Systems und dessen Modellrepräsentation, Fehler in Messdaten und die Anfälligkeit der Gesellschaft für klimatische und hydrologische Ereignisse. Solche Unsicherheiten kann man nicht unbedingt reduzieren oder quantifizieren. Sie zu untersuchen, kann aber dennoch zu einem verbesserten Verständnis des Klimawandels und dessen Folgen führen und damit auch die Entwicklung zukünftiger Modelle steuern sowie Entscheidungen bezüglich Anpassungsstrategien unterstützen. Nebst der Charakterisierung von Unsicherheiten nimmt auch der Aspekt der Robustheit eine zentrale Rolle ein: Robust sind Modelle, die auch unter veränderten Klimabedingungen zuverlässig sind. Robust sind Projektionen, die trotz Unsicherheit dieselbe signifikante Änderung aufzeigen und robust sind Anpassungsmassnahmen, die für die Gesellschaft unter verschiedenen Klimabedingungen hilfreich sind.

Die vorliegende Dissertation zeigt, dass Abflussprojektionen für die Schweiz trotz Unsicherheiten robuste Veränderungen aufzeigen. Dazu gehören eine Abnahme des Abflusses im Sommer sowie eine frühere Schnee- und Eisschmelze. Diese Ergebnisse basieren auf einem besonders breiten Ensemble von Modellketten, was zudem zeigt, dass eine Verminderung der projizierten Auswirkungen des Klimawandels auf den hydrologischen Kreislauf um ungefähr einen Faktor zwei möglich wäre, wenn die Treibhausgasemissionen stark reduziert werden. Die Verknüpfung von Klimamodellen und hydrologischen Modellen stellt eine grosse Herausforderung dar, insbesondere wegen systematischer Unterschiede zwischen beobachteten und simulierten Temperatur- und Niederschlagswerten in Klimamodellen. Diese Unterschiede - oft als Abweichung bezeichnet - können zu Anomalien in Abflusssimulationen führen. In dieser Arbeit wird gezeigt, dass Abweichungen auf regionaler Ebene (über der Schweiz) unter anderem entstehen, weil die Repräsentation der atmosphärischen Zirkulation auf synoptischer Ebene (über Europa) ungenau ist. Daher ist es schwierig, diese Abweichungen mit Nachbearbeitung (postprocessing) der Daten zu reduzieren. Zudem wird gezeigt, dass diese Abweichungen nicht nur von der Beschränktheit der Klimamodelle herrühren, sondern auch von Unsicherheiten in Beobachtungsdatensätzen sowie von der

natürlichen Variabilität des Klimasystems. Diese Ergebnisse unterstreichen wie wichtig es ist, die Ursachen systematischer Abweichungen von Klimamodellen zu identifizieren und zu berücksichtigen, um robustere Korrekturmethoden zu entwickeln. Zur Beurteilung der Robustheit hydrologischer Modelle unter sich ändernden Klimabedingungen, wird in dieser Arbeit eine neue Methode vorgeschlagen. Sie basiert auf dem Vergleich von simulierten und beobachteten Trends im Abfluss und in der Verdunstung während der letzten Jahrzehnte. Die Methode erlaubt den potenziellen Vorteil zeitabhängiger Modellparameterwerte im Vergleich zu einer Standardkonfiguration mit konstanten Werten zu untersuchen. Abschliessend wird in dieser Arbeit argumentiert, dass verbesserte Abflussprojektionen alleine nicht genügen, um signifikante Fortschritte in der Ausarbeitung von Adaptionsstrategien zu erlangen. Ein weiterer wichtiger Schritt ist die Förderung des interdisziplinären Dialogs zwischen Wissenschaftlern und Entscheidungsträgern. Zusätzlich, in Bezug auf die Lehre an der Universität, sollen die Studierenden auf Unsicherheiten inhärent zum Klimawandel sensibilisiert werden und Möglichkeiten mit diesen Unsicherheiten umzugehen aufgezeigt werden. Dieses Konzept wurde mit einem mehrtägigen Workshop an der Universität Zürich umgesetzt und die Hauptergebnisse sowie Empfehlungen für zukünftige, ähnliche Anlässe werden vorgestellt. Damit macht diese Dissertation deutlich, dass sowohl die Analyse von Stärken und Schwächen von Modellen als auch die Förderung des interdisziplinären Dialoges essentiell dafür sind, das Potenzial von Modellen im Hinblick auf den Klimawandel voll zu nutzen.

Résumé

Le climat est en train de changer, mais bien que l'augmentation de la température à l'échelle de la planète soit bien établie, ses impacts à l'échelle des bassins versants sont encore mal déterminés. Les impacts sur le débit des rivières sont d'une importance particulière pour la société, en raison du rôle des cours d'eau pour les ressources en eau, les écosystèmes, la production d'énergie, les loisirs, etc. L'analyse des changements hydrologiques induits par le réchauffement climatique repose principalement sur les modèles numériques, qui simulent l'évolution du climat et ses conséquences sur l'hydrologie des bassins versants. Cependant, comme les modèles sont, par nature, des représentations simplifiées de la réalité, il est crucial de prendre en compte leurs capacités et limites lorsque l'on interprète leurs simulations. Cela implique que les incertitudes des projections climatiques et de débit doivent être caractérisées. Ces incertitudes proviennent de différentes sources, telles que les futures émissions de gaz à effet de serre, la nature chaotique de l'atmosphère, notre compréhension du système et sa représentation dans les modèles, les erreurs dans les sets d'observations et les vulnérabilités de la société face aux événements climatiques et hydrologiques. Bien que ces incertitudes ne soient pas nécessairement réductibles ou quantifiables, leur exploration est capitale pour améliorer notre compréhension du changement climatique et de ses impacts, ainsi que pour guider le développement de futurs modèles et aider le choix de stratégies d'adaptation. En parallèle à la caractérisation des incertitudes liées au changement climatique, se développe une quête de plus de robustesse; avec des modèles robustes, fiables même dans des conditions climatique changeantes; avec des projections robustes qui sont en accord sur des changements significatifs malgré leurs incertitudes; avec des mesures d'adaptation robustes qui bénéficient à la société pour un large spectre de conditions climatiques.

Cette thèse montre que malgré les incertitudes des projections de débit en Suisse, des changements robustes émergent, tels que la diminution des débits en été et une pointe de débit liée à la fonte des neiges et des glaciers plus précoce. Ces résultats sont basés sur un ensemble de chaînes de modèles particulièrement large, qui révèle également que les impacts sur les régimes hydrologiques projetés pourraient être réduits d'environ un facteur deux, si des efforts conséquents de réduction des émissions de gaz à effet de serre étaient réalisés. La combinaison de modèles climatiques et hydrologiques est délicate, en particulier à causes de différences systématiques entre les simulations de précipitation et température fournies par les modèles climatiques et les observations de ces mêmes paramètres. Ces différences, couramment appelées 'biais', peuvent mener à des anomalies dans les simulations hydrologiques. Il est montré ici qu'une part de ces biais à l'échelle régionale (en Suisse) provient de la représentation de la circulation atmosphérique à l'échelle synoptique (au dessus de l'Europe), ce qui compromet notre capacité à réduire ces biais a posteriori. Il est également prouvé que ces biais ne proviennent pas uniquement des limitations des modèles climatiques, mais aussi d'incertitudes liées aux sets d'observations et à la variabilité naturelle du système climatique. Ces résultats soulignent l'importance d'identifier et de prendre en compte les causes de ces biais pour progresser vers des méthodes de correction plus robustes. Dans le contexte de la modélisation hydrologique, une nouvelle méthode pour explorer la robustesse des modèles hydrologiques appliqués dans des conditions climatiques changeantes est proposée. Cette méthode est basée sur l'évaluation de la représentation par les modèles des tendances de débit et d'évapotranspiration observées durant les dernières décennies. Elle permet d'évaluer les bénéfices potentiels d'une calibration hydrologique dont la valeur des paramètres varie dans le temps, par rapport à une configuration standard basée sur des valeurs constantes. Finalement, il est relevé que de meilleures projections ne suffiront pas à elles seules pour faire progresser le développement de stratégies d'adaptation de manière significative. Des étapes supplémentaires sont nécessaires, notamment la stimulation du dialogue interdisciplinaire entre scientifiques et décideurs, et, dans le cadre universitaire, la sensibilisation les étudiants aux incertitudes inhérentes au changement climatique et aux méthodes pour les gérer. Nous avons mis cette idée en pratique en organisant une rencontre de plusieurs jours à l'Université de Zurich. Nous présentons ses principaux résultats et des recommendations pour l'organisation de futures rencontres de ce type. Cette thèse conclut qu'explorer les capacités et limites des modèles, ainsi que favoriser le dialogue interdisciplinaire, est essentiel pour pleinement bénéficier du potentiel des modèles dans le contexte du changement climatique.

Contents

Chapter 1

Introduction

1.1 A peek into the future using models

If you check the current weather forecast when deciding whether to pack an umbrella before leaving home or where to spend the coming week-end, then your decision partially depends on the output of an atmospheric model. There is some uncertainty, the weather does not always turn out as in the forecast, yet we still rely on weather forecasts, because they are more informative than pure guesses. Now, what if we could better understand why models sometimes go wrong, would not that help to improve future forecasts? And if some part of the forecast could not be fixed, and would make the forecast particularly uncertain in some cases, would not it be great to better identify those situations, and then communicate this uncertainty? Finally, what if the weather forecast was tailored to your needs, and the information provided was really helping you to make a decision, while fully exploiting the potential of models?

This thesis focuses on the changes in river discharge in Switzerland until 2100. Although this is a longer time scale than that of weather forecasts, the research presented here deals with similar questions and addresses similar challenges to those outlined in the weather analogy of the previous paragraph. A common characteristic of weather forecasting and the investigation of future river discharge is that both rely on models to get a peek into the future.

To explore future climatic and hydrological conditions, the research community heavily relies on models. 'A model is a simplified representation of a system, with twofold purpose to enable reasoning within an idealized framework and to enable testable predictions of what might happen under new circumstances' (Gupta et al., 2012). From a technical perspective 'quantitative models are normally expressed as sets of assumptions and mathematical equations and implemented as computer codes' (Beven, 2008). These definitions are broad enough to apply to a wide range of models, in particular those investigated in this thesis, which simulate the climate system (climate models) and the hydrological processes relevant at the catchment scale (hydrological models). Models are simplified representations of the reality, hence are imperfect and subject to uncertainties. They nevertheless allow for the exploration of future conditions, and hence can be used for the anticipation and mitigation of future changes. They also enable the scientific community to test and improve its understanding of our environment, for instance of the climate system and catchment hydrology.

The increase of concentration of greenhouse gases in the atmosphere, as a result of human activities, is leading to an increase of the global temperature (IPCC, 2013). What will be the consequences for the water cycle, for instance in terms of the snow pack or flood frequency? The answer is dependent on a wide range of aspects, such as future greenhouse gas emissions, the response of the climate system and watersheds to these emissions, and the region of interest.

1.2 Modeling through spatial and temporal scales

Discharge in a catchment is the result of processes occurring in a wide range of spatial scales, including synoptic atmospheric flows, the effect of mountain ranges on precipitation, water infiltration in the pores of the soil and plant transpiration through their stomates. Similarly, we expect from models that they correctly capture transient changes spanning more than 100 years, and at the same time, that they simulate realistically events like thunderstorms and the following peak discharge, which occur over a period of about one day. Capturing and then combining these different scales is a considerable challenge, which is broken down to be addressed. The different scales are represented by different models (Figure 1.1), under the premise that processes are best represented at the scale at which they occur, which leads to different levels of representation of the spatial heterogeneity (spatial discretizations). The global scale is dealt with by general circulation models (GCMs) which are run over the entire planet, on a grid with a horizontal resolution of the order of ~200 km for the models used in this thesis. The regional scale is simulated by regional climate models (RCMs), which are run on a limited area, for instance Europe, at a resolution of ~25 km. Finally, catchments are discretized in units of similar characteristics, which in this thesis can be as small as 250x250 m^2. Each of these models represents smaller-scale (sub-grid) processes in a simplified way, by resorting to parameterizations.

Models are combined in an attempt to combine their strengths. For instance, only GCMs are able to simulate the influence of oceans on climate and to cover the whole globe, but they cannot properly account for the influence of the Alps because of their coarse resolution. This task is performed by RCMs, which are run at higher resolution over a limited area, a process known as dynamical downscaling (Fowler et al., 2007). RCMs are however still too coarse to capture changes in variables such as snow conditions for a given mid-elevation resort or the water input into a lake. RCM outputs are hence further downscaled to account for small scale features, for instance to reproduce local temperature profiles. They can then be used to force hydrological models and produce discharge projections.

The different elements of the model chain exert a different influence on the simulated discharge. GCM-RCMs tend to have a larger influence on the projected discharge than the hydrological models (e.g., Bosshard et al., 2013a). Further, models being nested into each other, with the larger scale model driving the smaller scale model, a large part of the changes and uncertainties at a small scale stems from a larger scale. van Ulden and van Oldenborgh (2006) illustrate for instance how errors in fields simulated by the GCM propagate down the chain, and impact surface parameters in RCM simulations, such as temperature and precipitation.

The combination of these different models running at different spatial scales is delicate, the interface between regional climate models and hydrological models being particularly challenging. Their combination is necessary for the production of discharge projections, yet the presence of biases (systematic differences between model simulations of observations) in climate models represents a major barrier, as biases can lead to severe anomalies in hydrological simulations. In order to reduce them, a key step is to investigate where they

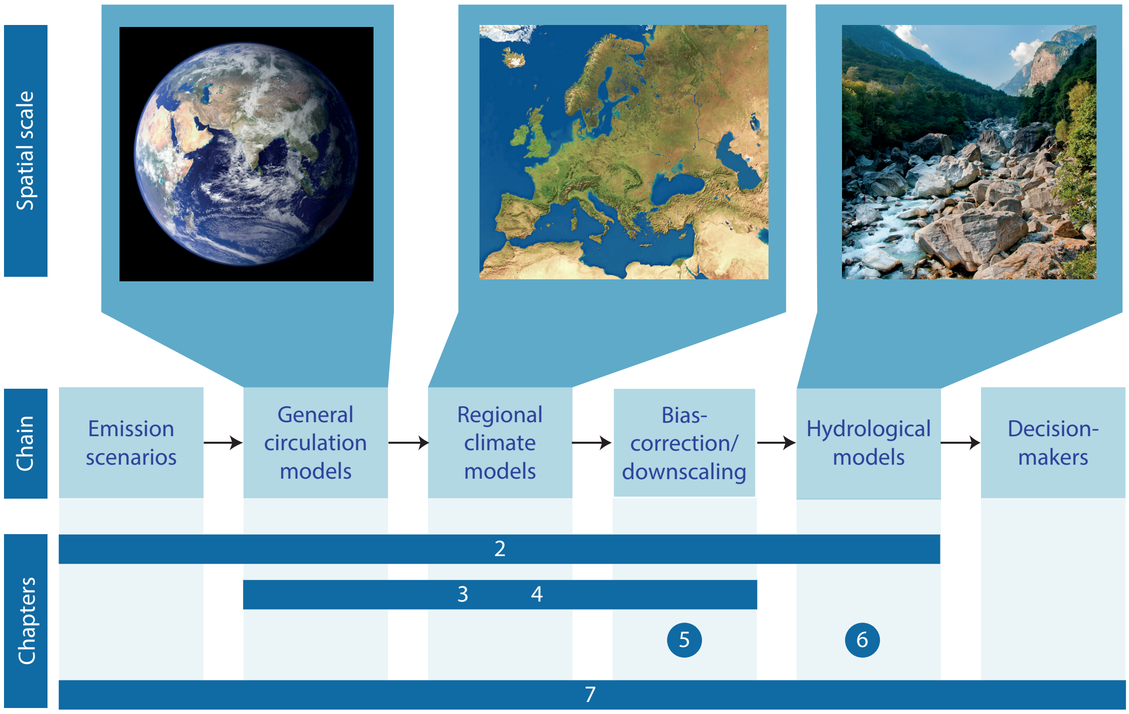

Figure 1.1: Schematic depiction of the spatial scales involved in impact modeling (top row), their representation in models (middle row) and the focus of this thesis' chapters (bottom row).

come from (Bellprat et al., 2013), how much they are influenced by natural variability and errors in observational datasets (Sunyer et al., 2013) and how these biases evolve under changing climate conditions (Buser et al., 2009; Chen et al., 2015).

1.3 Uncertainties and robustness in model simulations

Models are simplified representations of reality, which implies that their simulations are subject to uncertainties. There is a growing effort, both in climate science and hydrology, to better explore, characterize and communicate uncertainty (Mastrandrea et al., 2010; Pappenberger and Beven, 2006). This results in an ongoing shift from deterministic approaches, relying on 'models that with a set of initial and boundary conditions have only one possible outcome or set of predictions' (Beven, 2008), to probabilistic approaches, in which model uncertainty is embraced and usually represented by an ensemble of simulations and/or probabilities.

The characterization and communication of projection uncertainties are of high priority, since they enable a more transparent communication of model limitations, and if risk is expressed in terms of 'damage × probability of occurrence', uncertainty quantification provides a priori a more solid basis for quantitative risk assessment. To explore model uncertainty, the most widely adopted approach is to run different models and use the dispersion among their simulations as an uncertainty estimate. This approach can be applied to GCMs (e.g., Taylor et al., 2012), RCMs (e.g., Mearns et al., 2009; van der Linden and Mitchell, 2009), or chains of models involving a succession of GCMs, RCMs and hydrological models (e.g., Wilby and Harris, 2006; Bosshard et al., 2013a). Different methods can then be used to investigate projection uncertainty, such as analyses of variance (e.g., Hawkins and Sutton, 2010; Bosshard et al., 2013a) or Bayesian techniques (e.g., Tebaldi et al., 2005; Fischer et al., 2012). Similar setups allow for the investigation of specific sources of uncertainty, such as parameter uncertainty (Bellprat et al., 2012) or natural variability (Deser et al., 2012a).

The reduction of these uncertainties is a considerable challenge. Uncertainties cascade in the model chain from one model to the other, and result from processes interrelated in mostly non-linear ways and taking place in a wide range of spatial and temporal scales. Further, part of the uncertainty is related to the chaotic nature of the atmosphere, meaning that small differences in the model initial conditions can lead to significant differences in the model outcome. The natural climate variability of the climate system is an example of chaotic behaviour, and is an irreducible source of uncertainty (Deser et al., 2012a). Further, there are still considerable gaps in our knowledge of how the atmosphere (in particular its dynamics) will respond to global warming, and it will take time to fill those gaps (Shepherd, 2014). Increasing model complexity (often perceived as realism) may not suffice to reduce projection uncertainty, since it usually means that processes that used to be parameterized need to be explicitly resolved, which involves new uncertainties (Bierkens et al., 2015). Overall, model design and evaluation under present and future conditions is not exempt of uncertainties either, given the limitations of state-of-the-art observational data sets (Gómez-Navarro et al., 2012; Sunyer et al., 2013) and the absence of observations for the future (Refsgaard et al., 2013).

Because of the difficulty to reduce these uncertainties, new methods are being proposed, that enable the identification of significant changes despite uncertainties and decision-making in presence of uncertainties. Several of these methods rely on an increased robustness. Robustness can be understood in different ways. When models agree on a change, and when the average change is a significant change, then the projection may be considered robust, although the individual models for instance do not agree on the exact amplitude of the change (Knutti and Sedláček, 2012). Similarly, when a decision has to be made on the basis of uncertain projections, a robust decision is a decision that leads to acceptable results for a broad range of climate outcomes (Wilby and Dessai, 2010). Finally, a model or post-processing method is defined as robust when it performs well even when executed under climatic conditions significantly different from that of its calibration period (Coron et al., 2012).

1.4 Finding a balance between parsimony and realism

In a recent interview[1], Sten Bergström, the co-creator of the hydrological model HBV was asked: 'The HBV model has been and is still one of the most used hydrological models, what is the reason for its success?', to which he replied: 'I think that we found a sound balance between complexity and the information contained in most operational datasets'. This illustrates the importance of the correspondance between the level of sophistication of a model and the available resources, or in more general terms, the question to be addressed.

The model HBV, which is one of the hydrological models used in this thesis, relies on simple structure of linear reservoirs ('leaking buckets', Figure 1.2). Such models are termed conceptual, more complex models are usually called process-based, as some argue

[1] http://hepex.irstea.fr/the-hbv-model-40-years-and-counting/

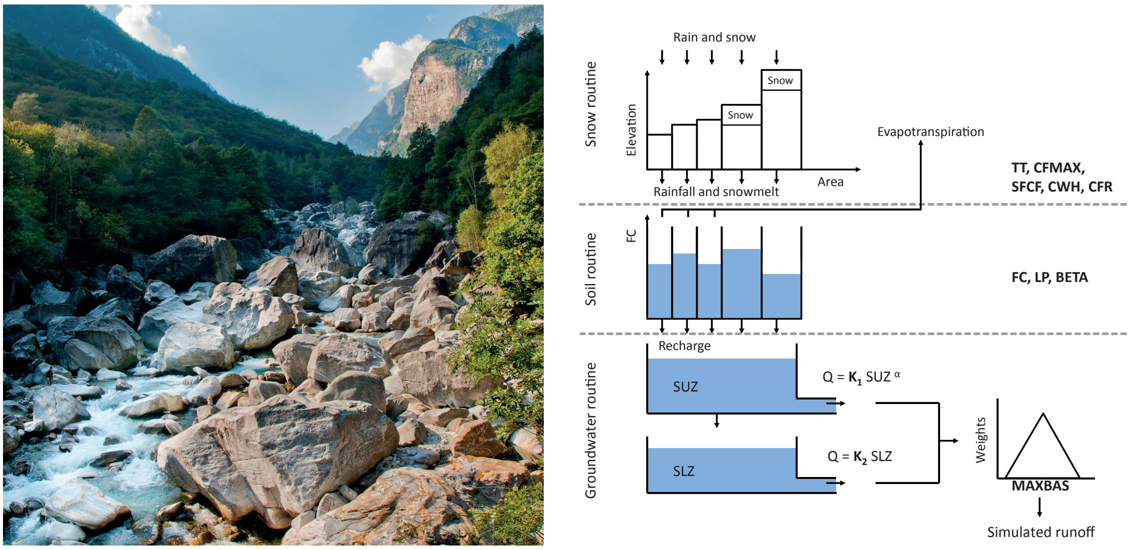

Figure 1.2: A catchment (left) and its representation in the conceptual model HBV (right, schema modified from Staudinger et al., 2015).

that no hydrological model is based on correct physical principles (Beven and Young, 2013). When comparing HBV structure to the image of a watershed (see Figure 1.2), one might doubt that such a simple model can reproduce the complex processes leading to discharge. Yet HBV and other models of a similar complexity (e.g., VIC model, Liang et al., 1994) are used in a large proportion of impacts studies. It is noteworthy these models have changed relatively little since their first publication, when at the same time there has been an uninterrupted increase of complexity (and decrease of resolution) in climate models (Knutti and Sedláček, 2012).

It does not mean that more complex hydrological models were not developed. Part of the hydrological modelling community strongly supports the increase of complexity, resolution and level of detail. Yet, using more complex formulations and structures implies that data requirements will be larger and computing times longer. An example is hydrological modeling at ‘hyperresolution’, i.e., on the order of 1 km over continental to global domains, which would involve massive computational resources to solve as many as 10^9 unknowns (Wood et al., 2011). In contrast, other scholars argue that there is only a limited number of parameters that can be determined based on the information available in most catchments (Jakeman and Hornberger, 1993; Perrin and Michel, 2001; Kirchner, 2006; Beven and Cloke, 2012). A similar dilemma can be witnessed to some extent for climate modelling. At one end of the spectrum are climate models at extremely high resolution (e.g., 2.2 km in Ban et al., 2014), so expensive to run that simulations are limited to a few decades under present climate and a few decades under future climate. There are few extreme events in such short periods and uncertainty sampling is particularly costly. At the other end of the spectrum, weather generators can be run for centuries on a laptop, yet it is unclear how well they capture future climate without the strong physical bases of climate models.

In summary, there is a trade-off to be found between increasing model complexity and exploring model space and uncertainty. A key challenge is to pinpoint the ‘appropriate model complexity’ for a given purpose (Gupta et al., 2012).

1.5 The calibration of hydrological models under climate change

There is not necessarily a clear link between hydrological model parameters and real world characteristics. As expressed by Savenije (2009), 'the dominant processes in the hydrological cycle only become apparent at larger scales, and these processes are not just the sum of processes occurring at micro-scales'. A consequence is that parameter values cannot be directly inferred from catchment characteristics, or measured in the field (Wagener, 2003). Instead they have to be determined by model calibration, which illustrates the challenge of efficiently transferring field observations into models (Seibert and McDonnell, 2002). Calibration is the process by which the values of model parameters are adjusted in order to better simulate the observed behavior of the catchment, usually summarized by its discharge. The correspondence between model simulations and observations is assessed quantitatively using a measure of skill ('objective function'), which is often the Nash Sutcliffe efficiency (Nash and Sutcliffe, 1970; Schaefli and Gupta, 2007), but can also be a multi-criteria metric. The search of the parameter space can be performed using different methods, such as Monte Carlo simulations (e.g., Finger et al., 2012), a genetic algorithm (e.g., Seibert, 2000) or Bayesian inference (e.g., Fenicia et al., 2014). Alternatively, parameter values can be inferred from other catchments of similar characteristics, a process known as 'regionalization' (e.g., Viviroli et al., 2009a). Such processes usually yield a number of parameter sets, which despite their differences, lead to model simulations of comparable skill, an effect called 'equifinality' (Beven, 2006).

As a complement to model calibration, models are then 'evaluated' (or 'validated' depending on the authors). This corresponds to the assessment of model's skill during a period different from the period used for the calibration. Model validation takes a new dimension in the context of modeling under a changing climate, because it is unclear how 'different' the climate during the calibration and the validation period should be for the validation to be meaningful, i.e., in order to trust the model's ability to deliver reliable results under a changing climate. Several approaches exist to assess the robustness of hydrological models under changing climate (Klemeš, 1986; Seibert, 2003; Coron et al., 2012; Refsgaard et al., 2013) and improving models' robustness is one of the grand challenges of hydrological research.

1.6 Decision-making under uncertainty

The need to better address and characterize uncertainties is now well established in the climate and hydrology communities. Yet the inclusion of these uncertainties into decision processes is arguably still in its infancy. In fact, although climate change research is often motivated by the support it can bring to decision-making, scientific knowledge by itself does not necessarily lead to better decisions.

A key issue is the mode of production of the projections. Uncertainties are usually propagated down the model chain (Figure 1.1) and are then communicated to decision-makers in charge for instance of designing adaptation strategies. This approach is refereed to as 'top-down' and is the dominant approach in impact studies. A key issue is that projections produced this way do not necessarily correspond to the needs of decision-makers (Dilling and Lemos, 2011). An example of top-down approach is the report 'Toward quantitative scenarios of climate change impacts in Switzerland' (CH2014 Impacts, 2014) to which I contributed as one of the two lead authors of the chapter 'Hydrological responses to climate change: river runoff and groundwater' (Rössler et al., 2014). This report was a unique opportunity to use a new set of climate projections (CH2011, 2011) and run a wide range of model chains in a systematic way, which provided new insights into the response of watersheds to climate change, into the sources of uncertainty in discharge projections and into robust changes in hydrological regimes. Yet we were confronted with an important shortcoming of the top-down approach, namely a lack of agreement between the projections provided and the data actually needed for impact modeling. The climate projections were delivered in three slices of 30 years, when glacier modeling requires transient projections. Further, the projections were derived using a delta change method, which does not capture well changes in variance. This precluded the study of extreme events relevant for many end-users. This example illustrates that for the potential of models to be exploited for impact modeling and the design of adaptation strategies, a good understanding of the needs and vulnerabilities of end-users is crucial. This premise is central to the development of alternative ways to deal with uncertainties, using 'bottom-up' approaches (Brown et al., 2012).

1.7 Outline and main research questions

The next six Chapters correspond to six peer-reviewed papers or manuscripts. They deal with different parts of the model chain (Figure 1.1) and focus on different research questions. Since these studies are always the result of collaborations, I use 'we' throughout this thesis.

Chapter 2: Robust changes and sources of uncertainty in the projected hydrological regimes of Swiss catchments. This study is the result of a coordinated effort to run a particularly wide range of model simulations to assess future discharge in Switzerland. It relies on three emission scenarios, 10-20 GCM-RCM simulations, two post-processing methods and three hydrological models combined factorially and applied to six catchments. The systematic approach and the experimental setup allowed us to address the following research questions:

1. Where does the uncertainty in discharge projections come from and how does this partitioning change with catchment characteristics and time?
2. How much does the complexity of hydrological models influence discharge projections?
3. Do robust changes in hydrological regimes emerge despite projection uncertainty?
4. Would limiting greenhouse gas emissions lead to significant reductions of the impacts on discharge?

Chapter 3: The influence of natural variability and interpolation errors on bias characterization in RCM simulations. Bias correction is performed using an observational data set as reference and in presence of natural climate variability. But before calling a bias a bias, it is important to understand which part of the differences between model simulations and the reference data set can actually be attributed to limitations of the climate models, and which part stems from natural variability and interpolation errors. The research questions were:

1. How much does natural variability influence the results of climate model evaluation over multi-decadal periods?

2. Does the unpredictable nature of internal variability preclude our chances to reduce climate model biases by post-processing?
3. How well can biases be characterized, and hence corrected, given the uncertainties in reference data sets?

Chapter 4: Propagation of biases in climate models from the synoptic to the regional scale: implications for bias-adjustment. A pre-requirement for bias-adjustment to be robust is that the reasons behind the biases are understood. Yet they are difficult to pinpoint, in some cases because processes leading to model errors are outside of the region considered for model evaluation. With this in mind, we explored how errors in the large-scale atmospheric circulation in the Alpine region can influence precipitation and temperature simulations in Switzerland. We assessed the circulation types simulated by 20 GCM-RCMs and explored the following research questions:

1. How well do GCM-RCMs capture the frequency and regime of circulation types under present climate?
2. How does this influence biases in GCM-RCM simulations of temperature and precipitation?
3. What are the implications for bias-adjustment and downscaling?

Chapter 5: Bias-correction for hydrological impact studies - beyond the daily perspective. An important question for the validation of a bias-correction method is whether it corrects for aspects that it was not trained to correct. This study focusses on a popular bias-correction method (quantile mapping) and explores the suitability of bias-corrected time series for the modeling of extreme hydrological events:

1. How well does quantile mapping based on daily values capture multi-day statistics relevant for flood and drought modeling?

Chapter 6: Trends in water balance as indicators of robustness for hydrological models in a changing climate. When hydrological processes change as a result of global warming, an important question is whether the value of parameters in hydrological models should be modified to capture these changes. We analysed observed and simulated trends in the water balance to explore this question and to investigate the robustness of the hydrological model HBV in four research catchments. We addressed the following research questions:

1. What are the trends in temperature, precipitation, discharge and evapotranspiration in the study catchments over 1971-2010?
2. How well are trends in discharge and evapotranspiration captured by HBV using external forcings alone, i.e., with no change in parameter values?
3. When HBV is calibrated over 10-year periods, is there a drift in parameter values and if yes, can this drift be related to observed changes in physical processes or to calibration artefacts?

Chapter 7: From products to processes: Academic events to foster interdisciplinary and iterative dialogue in a changing climate. Although assessing the projection uncertainty is an established field of research, it remains challenging to incorporate this uncertainty into decision-making processes. We believe that a promising way to advance with this challenge is to foster iterative and interdisciplinary exchanges between climate researchers and decision-makers, and to introduce students to uncertainties related to climate change at an early stage of their career. We therefore organized a workshop at the University of Zurich, to put this idea into test and explore the following aspects:

1. What are the dominant ways to deal with uncertainty in decision-making processes?
2. How do participants perceive the collaboration between researchers and decision-makers, and how does this perception evolve through the workshop?
3. To which extent can academic events contribute to foster iterative and interdisciplinary exchanges?

Chapter 2

Robust changes and sources of uncertainty in the projected hydrological regimes of Swiss catchments

Nans Addor[1], Ole Rössler[2], Nina Köplin[2,3],
Matthias Huss[4], Rolf Weingartner[2], and Jan Seibert[1,5]

[1] Department of Geography, University of Zurich, Zurich, Switzerland
[2] Oeschger Centre for Climate Change Research and Department of Geography, University of Bern, Bern, Switzerland
[3] Now at Swedish Meteorological and Hydrological Institute, Norrköping, Sweden
[4] Department of Geosciences, University of Fribourg, Fribourg, Switzerland
[5] Department of Earth Sciences, Uppsala University, Uppsala, Sweden

Published in Water Resources Research, 2014
doi:10.1002/2014WR015549

Abstract

Projections of discharge are key for future water resources management. These projections are subject to uncertainties, which are difficult to handle in the decision process on adaptation strategies. Uncertainties arise from different sources such as the emission scenarios, the climate models and their postprocessing, the hydrological models, and the natural variability. Here we present a detailed and quantitative uncertainty assessment, based on recent climate scenarios for Switzerland (CH2011 data set) and covering catchments representative for midlatitude alpine areas. This study relies on a particularly wide range of discharge projections resulting from the factorial combination of 3 emission scenarios, 10-20 regional climate models, 2 postprocessing methods, and 3 hydrological models of different complexity. This enabled us to decompose the uncertainty in the ensemble of projections using analyses of variance (ANOVA). We applied the same modeling setup to six catchments to assess the influence of catchment characteristics on the projected streamflow, and focused on changes in the annual discharge cycle. The uncertainties captured

by our setup originate mainly from the climate models and natural climate variability, but the choice of emission scenario plays a large role by the end of the 21st century. The contribution of the hydrological models to the projection uncertainty varied strongly with catchment elevation. The discharge changes were compared to the estimated natural decadal variability, which revealed that a climate change signal emerges even under the lowest emission scenario (RCP2.6) by the end of the century. Limiting emissions to RCP2.6 levels would nevertheless reduce the largest regime changes by the end of the century by approximately a factor of two, in comparison to impacts projected for the high emission scenario SRES A2. We finally show that robust regime changes emerge despite the projection uncertainty. These changes are significant and are consistent across a wide range of scenarios and catchments. We propose their identification as a way to aid decision making under uncertainty.

2.1 Introduction

Designing adaptation strategies to a changing climate implies making decisions on the basis of an uncertain knowledge of future conditions. Uncertainties in climate projections are being investigated on the basis of consequent coordinated experiments such as the global Coupled Model Intercomparison Projects (CMIPs, Taylor et al., 2012) or the European ENSEMBLES project (van der Linden and Mitchell, 2009). A clearly defined setup, with an agreement on, for instance, emission scenarios or the simulated region, makes simulations comparable and allows for detailed uncertainty assessments (Fischer et al., 2012; Knutti and Sedláček, 2012).

The main sources of uncertainties investigated in such studies are the emission scenario, the model formulation and parameterization, and the natural climate variability (Hawkins and Sutton, 2010). Studies exploring the future consequences of climate change on catchment discharge deal with additional sources of uncertainty, namely, the hydrological models and the downscaling or bias correction method. A key challenge is then to quantify the contribution of each of these sources to the uncertainty of the discharge projections (Bosshard et al., 2013a).

A common approach is to produce simulations based on a model chain in which a set of climate projections forces various impact models. Each element of the chain influences the simulated impact and thereby contributes to its uncertainty. Uncertainty propagation is then explored by applying an ensemble approach, i.e., by varying the elements of the chain and by interpreting the resulting projection variability as uncertainty.

This method was successfully applied in a number of cases (e.g., Horton et al., 2006; Wilby and Harris, 2006; Prudhomme and Davies, 2009; Dobler et al., 2012; Finger et al., 2012; Bosshard et al., 2013a). These studies and others constitute a solid basis for the exploration of uncertainty in discharge projections. They were conducted in a variety of locations, using different climate and hydrological models, and considered a wide breadth of hydrological parameters. This diversity provides us with a wealth of in-depth analyses but, at the same time, makes a comparison of results difficult. While there seems to be a general agreement on the dominant contribution of climate models to the uncertainty in discharge projections, different conclusions were drawn about the contribution of the hydrological models, for example. Because of the diversity of the setups employed, the causes for these differences are hard to identify. It is in particular unclear whether these results reflect more the differences between the model chains or between the study basins. Further, when studies focus on one or two catchments, the generalization of their results to other areas and the assessment of their robustness is difficult (Gupta et al., 2014).

Our study brings together three research groups applying their models to the same six catchments and relying on the same set of climate projections as input. Using the same setup for all catchments allows for the separation of the influence of the catchments from that of the other elements of the model chain. Another novel aspect is the use of climate projections under an intervention scenario (RCP2.6). Whereas most studies deal with one single scenario, or with scenarios involving no climate policy intervention, we investigate the impacts under the scenario RCP2.6 that implies stringent efforts to reduce emissions with the objective to keep global temperature increase below 2°C. We hence extend previous uncertainty studies by including, in a systematic way, a larger set of emission scenarios and by covering a wider range of catchment types. In this paper, the following questions are addressed: where does the uncertainty in discharge projections come from and how does this partitioning change with catchment characteristics? Do robust changes in regime emerge despite projection uncertainty? Would limiting emissions to RCP2.6 levels lead to significant reductions of the impacts on discharge?

2.2 Data and methods

2.2.1 Experimental design

We combined three emission scenarios, three regional climate model estimates (based on 10-20 climate model runs), two postprocessing methods, and three hydrological models in a factorial way, leading to a total of 54 model chains applied to six catchments and three future periods, 2020-2049, 2045-2079, and 2070-2099 (Figure 2.1). In total, 972 hydrological projections of 30 years each were produced and analyzed. The factorial design of this modeling experiment reduces the risk of sampling artifacts, i.e., of reporting a result which only stems from a particular combination of models and is not representative of the other possible combinations. Further, it allows us to disentangle the contribution of the different elements of the model chain to the ensemble variance and to quantify the eventual interactions between these factors.

2.2.2 Catchments

Discharge was simulated in six mesoscale catchments in Switzerland, which are representative of the main discharge types in midlatitude alpine regions. These basins were shown to react differently to climate change by Köplin et al. (2012), who clustered 186 Swiss mesoscale catchments into seven response types (Figure 2.2). Their mean elevation covers the range 700-2370 masl., their size varies between 231 and 1696 km^2, and two of them are partially glacierized. They can be characterized as humid, as their evaporation is mainly energy limited. The Venoge River flows on the plateau at the foothills of the Jura, a comparatively low-elevation mountain chain. Emme and Thur basins have their headwaters in the pre-Alps, whereas Rhone and Vorderrhein are alpine catchments, dominated by snow and icemelt. The Verzasca basin is located on the southern side of the Alps. Further information on the catchment characteristics are provided in Table 2.1.

Dams are present in most of the large Swiss alpine catchments. According to the Swiss Committee on Dams (www.swissdams.ch), three dams are upstream of the Brig gauging station for the Rhone catchment and seven are upstream of Ilanz for the Vorderrhein catchment. The reservoirs are in average larger in Vorderrhein catchment. Their total volume is about 31 $\times$ 10^6 m^3 for the Rhone and 260 $\times$ 10^6 m^3 for Vorderrhein, which

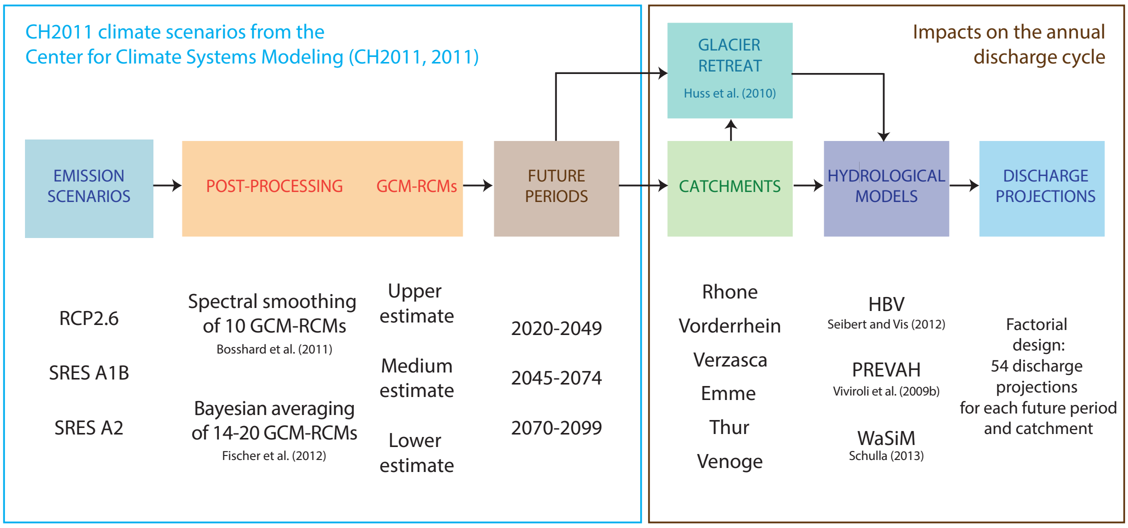

Figure 2.1: Flowchart of the experimental design. The boxes represent the model chain elements. For each element, several methods were used which are listed under the boxes and described in the main text. The three GCM-RCM estimates were derived from 10 to 20 different GCM-RCMs (see section 2.2.3).

corresponds to 2.3% and 24.9%, respectively, of the mean annual discharge at the gauging station for these two basins. The other four study basins, located at lower elevation, are essentially unregulated.

2.2.3 Climate projections

We used climate projections of the CH2011 data set (CH2011, 2011) from the Center for Climate Systems Modeling (C2SM). This recently released data set consists of two types of projections for Switzerland, both based on the climate model runs of the ENSEMBLES project (van der Linden and Mitchell, 2009) and both relying on the delta change approach. According to this technique, projections are produced by combining observations to additive (multiplicative) factors for temperature (precipitation). These factors were assessed in a deterministic way by spectral smoothing of 10 general circulation and regional climate model (GCM-RCMs) chains, and in a probabilistic way by a Bayesian multimodel approach combining 20 runs until 2050, and then 14 runs until 2099, respectively. CH2011 projections are not transient but available as mean temperature and precipitation changes over 2020-2049, 2045-2074 and 2070-2099, 1980-2009 being the reference period. This defines the periods over which hydrological simulations were run. The delta change factors were estimated from GCM-RCM simulations run with a daily time step. These factors were provided for each day of the year (365 values) and, once combined with station observations, yielded projections at the station scale. The interpolation within the catchments was performed by the hydrological models using the methods listed in Table 2.2. As the delta change method does not capture changes in variance, we restricted our analysis to the changes in the annual discharge cycle and did not consider extremes.

The production of the projections by the CH2011 team and their use to force hydrological models in this study can be summarized as follows. For the deterministic data set, the delta change factors were determined for each of the 10 GCM-RCM chains by spectral

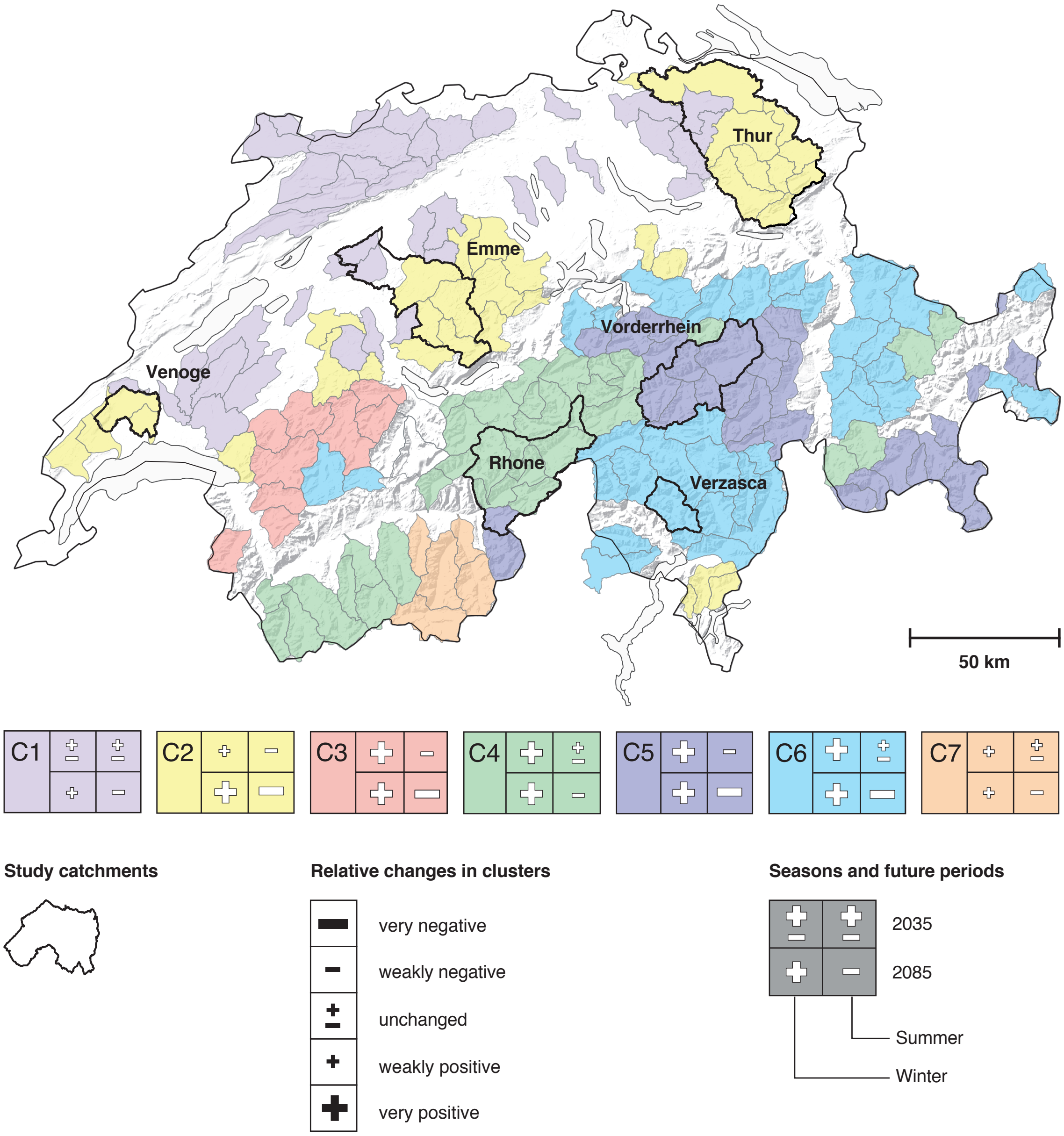

Figure 2.2: Map of the six study catchments within Switzerland. The color coding refers to clusters of catchments subject to a similar hydro-climatological changes. The discharge changes projected for these watersheds are summarized by the plus and minus signs. The clustering methodology and the model chain used to produce those projections are detailed in Köplin et al. (2012).

River Gauging station	Area[b] [km^2]	Mean elevation[b] [masl]	Elevation range[c] [masl]	Glacierized area in 1995[d] [%]	Mean annual T[e] [°C]	Mean annual P[e] [mm]	Mean annual Q[b] [mm]	Mean date of half flow computed using observed discharge[b]	Cluster of the hydro-climatological change signal[f]
Rhone Brig	913	2370	667 - 4256	22.3	0.1	1859	1480	6 July	C4
Vorderrhein Ilanz	776	2020	693 - 3609	2.9	2.1	1531	1347	3 June	C5
Verzasca Lavertezzo	186	1672	490 - 2856	0	4.8	2026	1763[g]	11 May[g]	C6
Emme Wiler	939	860	458 - 2216	0	7.3	1432	648	8 April	C2 - C1
Thur Andelfingen	1696	770	356 - 2500	0	7.7	1421	892	10 April	C2 - C1
Venoge Ecublens	231	700	383 - 1677	0	9.6	1216	577	27 February	C2

Table 2.1: Key characteristics of the six catchments[a]. [a]Mean annual values computed over the period 1980–2009, unless indicated otherwise. [b]Data source: Swiss Federal Office for the Environment. [c]Data source: 25 m digital elevation model of the Swiss Federal Office for Topography. [d]Data source: Fischer et al. (2014). [e]Data source: Swiss Federal Office of Meteorology and Climatology 2 km gridded data sets TabsD (Frei, 2013) and RhiresD (Frei and Schär, 1998; Schwarb, 2000; Frei et al., 2006, section 4.1). [f]Data source: Köplin et al. (2012). [g]Mean over 1990–2009.

smoothing. This technique enables the isolation of the precipitation and temperature annual cycles by removing fluctuations arising from natural variability, which may result in artifacts in the estimated climate change signal (see Bosshard et al., 2011, for more details). Hydrological simulations were run for each of the 10 GCM-RCM chains and then the lower (minimum of the 10), medium (median), and upper (maximum) estimates of discharge changes were derived. For the probabilistic data set, the 14-20 GCM-RCM chains were combined into probability distributions using a Bayesian multimodel combination algorithm. As this algorithm requires data to be normally distributed, precipitation data were transformed to their square root before their processing, and retransformed after it. Similarly, the internal decadal variability was subtracted before the multimodel combination and was then readded. The delta change factors corresponding to the 2.5, 50, and 97.5% quantiles were then extracted from the posterior distributions, after data retransformation and the readdition of the internal variability (see Fischer et al., 2012; Zubler et al., 2014, for more details on the method and on its application to the alpine range, respectively). These projections, corresponding to the lower, medium, and upper climate estimates, were used to force the hydrological models. Hence, in the deterministic case, each discharge simulation was driven by a single GCM-RCM chain, whereas in the probabilistic case, it was driven by, e.g., the median of the GCM-RCM chains. The name of the climate models used and of the institutions responsible for the simulations are provided in Bosshard et al. (2011, Table 1) and Fischer et al. (2012, Figure 1).

The CH2011 projections were produced for three possible evolutions of greenhouse gas emissions and concentration: SRES A1B, SRES A2, and RCP2.6. The first two storylines were defined in the Special Report on Emissions Scenarios (Nakicenovic and Swart, 2000, SRES) and used in the framework of CMIP3. The third storyline is the Representative Concentration Pathway RCP2.6 (Meinshausen et al., 2011) used for the more recent CMIP5 (Taylor et al., 2012). A1B is a midrange emission scenario that was chosen for the ENSEMBLES simulations. To gain insights into lower and higher emissions pathways without the computational cost of rerunning the climate models under these scenarios, the CH2011 team used pattern scaling to generate projections under the climate stabilization scenario RCP2.6 and under the more extreme A2 scenario (CH2011, 2011; Fischer et al., 2012). This technique is a statistical emulator relying on the principal assumption that any regional change is related to the global mean temperature signal in a linear way. Pattern scaling was applied to generate both precipitation and temperature projections at the regional scale, using scaling factors derived from temperature changes at the global scale (Fischer et al., 2012, Equation 12). The accuracy of the pattern-scaled projections primarily depends on the validity of the underlying assumption of linear relationship. Tebaldi and Arblaster (2014) concluded that it is broadly valid for CMIP3 and CMIP5 models and report geographical patterns of change consistent across different emission scenarios. They also review limitations of the method and emphasize in particular that pattern-scaled projections are more reliable for temperature than for precipitation, at large than at small spatial scales and are likely less reliable in presence of feedbacks or when the focus is on extreme events. In this study, in the absence of systematic GCM-RCM runs under A2 and RCP2.6, we consider the use of pattern-scaled projections as an adequate way to perform a first assessment of sensitivities in hydrological simulations to greenhouse gas emission scenarios.

To summarize, the projections under A1B were obtained from GCM-RCM simulations run under this scenario. In contrast, the projections under A2 and RCP2.6 do not rely on climate model runs under these scenarios, but were obtained using pattern scaling. Further, while A1B and A2 are nonintervention scenarios, considerable efforts are deployed under RCP2.6 to mitigate emissions. It follows that under this scenario, a global temperature increase by the end of the 21st century higher than 2°C relative to 1850-1900 is considered unlikely (i.e., 0-33% probability, IPCC, 2013). Note that in the CH2011 data set, RCP2.6 is referred to as RCP3PD, these two names corresponding to the same scenario (Meinshausen et al., 2011). We opted for RCP2.6, as it is named in the IPCC Fifth Assessment Report (IPCC, 2013).

The CH2011 setup does not allow for the separation of the GCM-RCM uncertainty into climate model uncertainty and natural climate variability. Climate model uncertainty is potentially reducible via a better understanding of global warming processes and models improvement, whereas natural climate variability is a stochastic and intrinsic part of the climate system, and is irreducible. Characterizing the future natural climate variability could for instance be achieved using projections from the same climate model started from different initial conditions (Deser et al., 2012a) or by employing a weather generator constrained by GCM-RCM projections (Fatichi et al., 2014).

2.2.4 Hydrological modeling

Three hydrological models were used to evaluate the sensitivity of discharge projections to model structure, resolution, and calibration: HBV (Seibert and Vis, 2012), PREVAH (Viviroli et al., 2009b), and WaSiM (Schulla, 2013). HBV and PREVAH are semidistributed (HBV was applied using 100 m elevation zones and PREVAH using hydrological response units (HRUs)), whereas WaSiM is run on a 250 × 250 m^2 grid. HBV and PRE-

VAH are based on a similar reservoir structure, while WaSiM uses a more process-oriented approach. HBV and WaSiM were calibrated, whereas PREVAH parameter values were obtained by regionalization (Viviroli et al., 2009a). PREVAH simulations hence represent natural discharge conditions even in catchments with major reservoirs. The hydrological models are contrasted in more details in Tables 2.2 and 2.3 and their coupling to the glacier model is discussed in section 2.2.5.

The reservoirs and the dam operations were not explicitly implemented in the hydrological models. Although their implementation might have enabled a better reproduction of the observed discharge under current climate, considerable uncertainties exist about the evolution of energy consumption and about the importance of hydropower production in comparison to other energy sources by the end of the century. One way to account for these uncertainties would be to formulate scenarios depicting different socioeconomic developments at the Swiss and the international level, possibly in a similar fashion as the SRES emission scenarios (Nakicenovic and Swart, 2000). In this study, however, we restrain our attention to the uncertainties stemming from the future emissions and the models used to produce hydrological projections, and assume that the simulated discharge corresponds to natural discharge. This assumption is realistic in the Rhone catchment, where the total reservoir volume corresponds to only 2.3% of the annual discharge, but is less robust in the Vorderrhein catchment, which is more heavily managed (24.9%). In the four other catchments, the discharge is not significantly affected.

Initial results indicated that the performance of the hydrological models under present climate was overall good and that they successfully capture the main features of the annual discharge cycles (Figure 2.3). Note that themodels performance in the Vorderrhein catchment is slightly impeded by discharge perturbations induced by hydropower production. Differences between simulated and observed annual cycle over the reference period (1980-2009) are assumed to be constant in the future. The overall approach relies on the assumption of time stationarity of the biases in the climate and hydrological simulations, although evidence for nonstationaries exist (Maraun, 2012; Merz et al., 2011). We hence defined changes in discharge as the differences between the projected discharge and the discharge simulated using 1980-2009 meteorological observations. We considered the annual cycle as captured by monthly averages and computed seasonal averages for summer (June, July, and August, referred to as JJA in continuation) and winter (December, January, and February, DJF). Further, we investigated the earlier occurrence of peak river runoff in spring to summer, one of the most prominent discharge changes in a warmer climate in snow-dominated regions. This shift in seasonality is generally attributed to two main factors: the higher proportion of precipitation falling as rain instead of snow, and earlier snowmelt and consecutive glacier melt (e.g., Barnett et al., 2005). To consider the combined effect of these two processes using one metric relying on discharge data, we considered the half-flow date (HFD) introduced by Court (1962), i.e., the date on which the cumulative discharge since the beginning of the hydrological year (starting on 1 October) exceeds half of the total annual discharge. This date is often preferred to that of the maximum daily discharge, which can result from a punctual precipitation event. The HFD has been broadly used to investigate the effects of climate change on streamflow timing (e.g., Cortés et al., 2011; Stewart et al., 2005).

Although Whitfield (2013) recently pointed out that the HFD is not a reliable indicator of the timing of snowmelt, we use it here from a more general perspective to assess the shift from snow-dominated regimes (nival) to more precipitation-dominated regimes (pluvial). In other words, we use the HFD as a regime metric, with a later HFD corresponding to

	HBV	PREVAH	WaSiM
Time step	Daily	Daily	Daily
Spatial units	100m-elevation zones (average area: 33 km^2; range for all catchments: 8 to 77 km^2)	Hydrological response units (HRUs) based on a 500 × 500 m^2 grid (average HRU area: 2.4 km^2; range for all catchments: 0.6 to 5.6 km^2)	Distributed on a 250 × 250 m^2 grid
Number of land cover classes	1	22	25
Snow and glacier melt parametrization	Snow and glacier melt: degree-day approach with aspect correction (Konz and Seibert, 2010)	Snow and glacier melt: temperature-index approach under consideration of the daily potential direct radiation (Hock, 1999)	Snow melt: degree-day approach with aspect correction. Glacier melt: temperature-index approach under consideration of the daily observed direct radiation (Schulla, 2013; Hock, 1999)
Evapotranspiration parametrization	Potential evapotranspiration: Oudin et al. (2005)	Actual evapotranspiration: Penman-Monteith	Actual evapotranspiration: Penman-Monteith
Input variables	Precipitation and temperature	Precipitation, temperature, sunshine duration, radiation, relative humidity, wind speed	Precipitation, temperature, radiation, relative humidity, wind speed
Interpolation of the input variables within the catchment	Projections combined using Thiessen polygons and then adjusted to each elevation zone using lapse rate for temperature and a similar relative factor for precipitation	Projections interpolated using detrended inverse distance weighting	Projections interpolated using inverse distance weighting and elevation-based regressions
Calibration or regionalization method	Calibration using a genetic algorithm (Seibert, 2000)	Calibration using an iterative search algorithm (Viviroli et al., 2009b) and subsequent regionalization (Viviroli et al., 2009a)	Two-step procedure: first manually, than coupling WaSiM with the parameter estimation tool PEST (Doherty, 2005)
Calibration period	1982-1995, except for the Verzasca (1992-1999)	1994-1997	1993-1998
Evaluation period		2000-2009, see Figure 2.3	
Average computing time for a 30-year simulation of one catchment	∼1 sec on a single computer core	∼4 min on a single computer core	∼20 min on a 32-core machine

Table 2.2: Key characteristics of the three hydrological models as used in this study.

	HBV	PREVAH	WaSiM
Main advantages	Low data requirements, simple structure allowing for fast calibration and execution, and facilitating the assessment of parameters influence on the simulated discharge	Computationally efficient (due to HRUs representing soil-land cover combinations), physical description of evapotranspiration (Penman-Monteith)	Detailed description of the catchment spatial heterogeneity allowing for the investigation of hydrological processes at a comparatively small scale
Main limitations	Coarse representation of the catchment spatial heterogeneity, empirical parametrization of snow and ice melt and of evapotranspiration	HRU structure inflexible, comparatively high data demand when Penman-Monteith is applied, hydrological processes in small catchments (<10 km^2) not well represented	High data requirements and computation time, the high spatial disaggregation makes parameters identification challenging
References for more detailed information on the models	Lindström et al. (1997); Seibert and Vis (2012)	Viviroli et al. (2009b); Zappa and Gurtz (2003)	Schulla (2013)
Other examples of application	Teaching (Seibert and Vis, 2012), simulation of design floods (Harlin and Kung, 1992), hydrological change detection (Gebrehiwot et al., 2013)	Parameter regionalization and flood estimation in ungauged catchments (Viviroli et al., 2009a), operational discharge forecasting (Zappa et al., 2008; Addor et al., 2011)	Investigation of the effects of land use changes on flood generation (Merta et al., 2008)and of the impacts of climate change on groundwater recharge (Neukum and Azzam, 2012) and on soil moisture (Rössler et al., 2012)

Table 2.3: Table 2.2 continued: key characteristics of the three hydrological models as used in this study.

higher elevation and more snow-dominated regimes. This relation appears clearly from the HFD computed using observed discharge in the six catchments (Table 2.1).

2.2.5 Modeling of glacier retreat and contribution to discharge

Ice melt significantly modifies the hydrological regime of glacierized catchments as it acts as a source of runoff after the disappearance of the snow cover (Barnett et al., 2005). The representation of glaciers in hydrological models is important to realistically simulate discharge during the summer months, and changes therein with respect to atmospheric warming (e.g., Horton et al., 2006). The hydrological models applied in this study include modules for glacier melt but do not describe the more complex processes of dynamic changes in glacier thickness and size over time. Therefore, glacier melt was simulated in two steps.

Glacier retreat was first modeled under the different climate projections and the resulting extent was provided as input to the three hydrological models. The contribution of glacier melt to discharge was then simulated in a second step by the hydrological models, i.e., not by the glacier model itself. Below, we provide more details and references about the modeling of glacier retreat and of the contribution of glacier melt to discharge.

The changes in future glacier coverage in the catchments were assessed by combining different glaciological models at high spatial resolution. The present glacier ice thickness

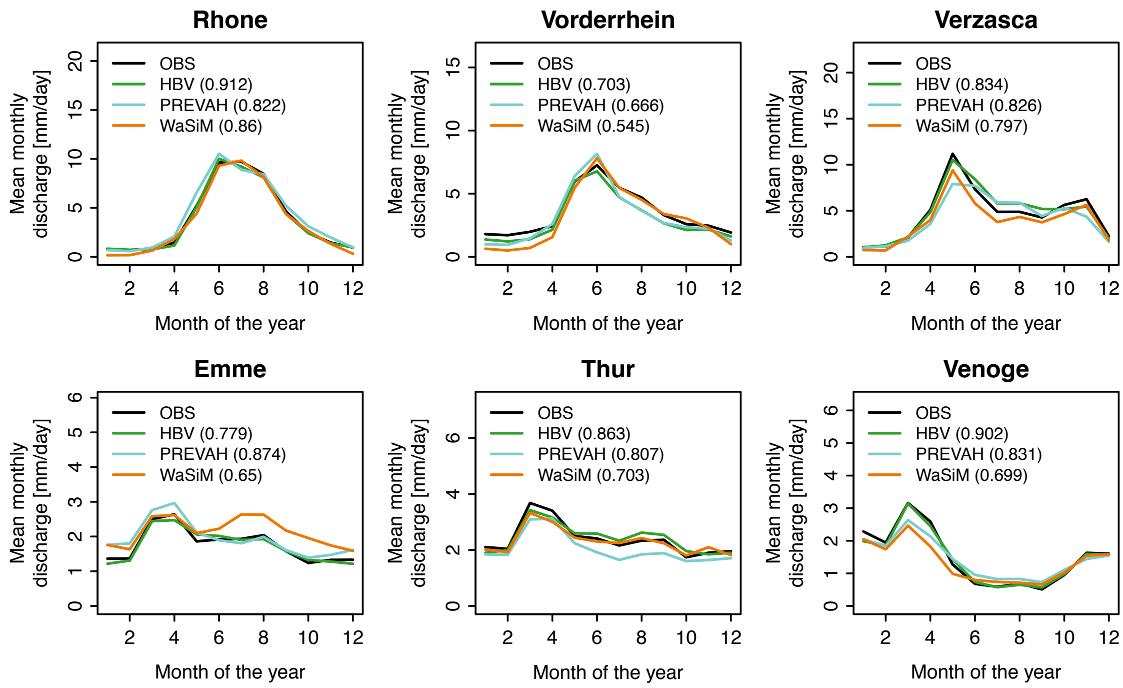

Figure 2.3: Performance of the three hydrological models under present conditions. Annual cycle observed (black) and simulated by the three hydrological models for the evaluation period 2000-2009. The Nash-Sutcliffe efficiency (Nash and Sutcliffe, 1970) for each model is given in brackets following its name.

distribution exerting an important control on the rate of glacier retreat and volume loss was evaluated. We used a model (Huss and Farinotti, 2012) to invert local ice thickness from glacier surface topography using the principles of ice flow dynamics. Glacier extent for the year 2010 was obtained from an inventory of all glacierized surfaces (Fischer et al., 2014). Surface mass balances and 3-D glacier geometry changes were calculated using a detailed glacier model (Huss et al., 2010b). This model was run at daily resolution on a 25 m grid and takes into account snow accumulation distribution, the influence of radiation on ice melting according to (Hock, 1999) and calculated glacier retreat based on a mass-conservation approach. The glaciological model was calibrated with a variety of field data covering the entire 20th century (Huss et al., 2010a). Annual mass balances from 50 glaciers with detailed results were then extrapolated to every single glacier in the catchment by applying a multiple regression approach. Thus, glacier-specific transient annual series of the glacier mass budget were obtained and used to drive the model for every glacier (152 in the Rhone catchment, 105 in the Vorderrhein catchment). The model was validated over the period 1973-2010 for which the change in glacier extent is known. The overall change in ice-covered area is captured within 5%, and the retreat rates of the glacier termini are reproduced in general although some differences for individual sites are evident (Figure 2.4). Finally, the glacier model was forced until 2100 with the same climate scenarios as the three hydrological models. We also applied the same approach to downscale the climate data to individual glaciers. From the transient model runs, we then extracted glacier ice coverage for 2035, 2060, and 2085 and assumed it to be constant over each of the 30 year period.

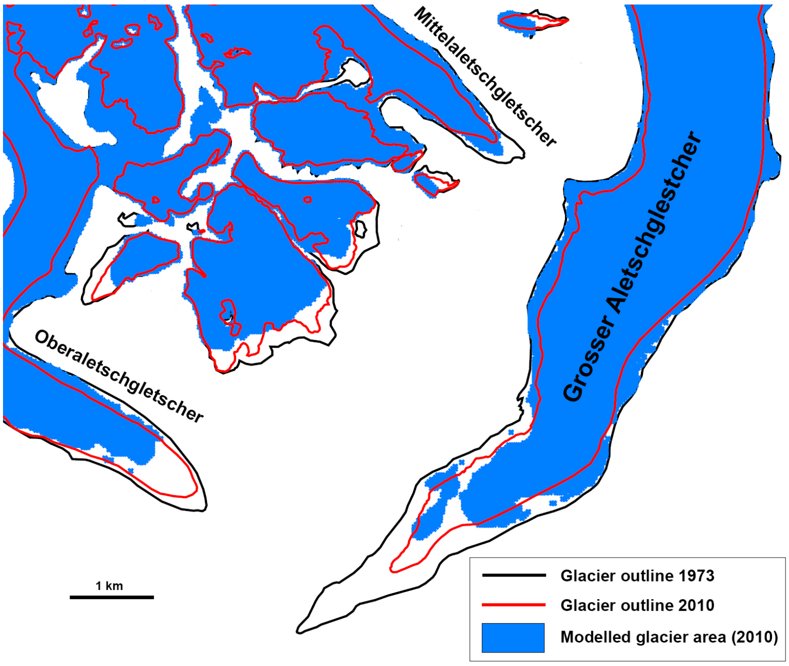

Figure 2.4: Validation of the glacier retreat model for the termini of the largest glaciers in the Rhone catchment. The model was initialized with the observed glacierized area in 1973 (black line) and run until the year 2010 when glacier extent is known (observed glacial extent as red line, simulated glacial extent as blue area).

These glacier simulations were used as input for the three hydrological models. Again, the prescribed glacier extent is the same in each case, but the melt and the resulting contribution to discharge were computed separately by each model. HBV relies on a degree-day approach considering temperature as the only driver of snow and ice melt. In the version used here, two multiplicative factors were added to the basic formulation of the degree-day approach in order to reflect the influence of exposure on snow and ice melt and to account for the lower albedo on ice than on snow (Konz and Seibert, 2010). Meltwater is transferred to the groundwater reservoirs, the outflow of which being then routed to the catchment outlet using a triangular weighting function. PREVAH and WaSiM are principally based on the temperature-index method developed by (Hock, 1999), which is an extension of the degree-day method. It involves melt factors for snow and ice, and importantly, it includes radiation coefficients for snow and ice, and accounts for direct solar radiation. Meltwater is routed to the underlying catchment using three linear reservoirs with different storage coefficients, for ice, firn, and snow, respectively.

2.2.6 ANOVA

We chose the projection variance as an estimate of their uncertainty and used an analysis of variance (ANOVA) technique to quantify the contribution of the different sources of uncertainty to the final uncertainty. The uncertainty partitioning relied on the following model

$$\Delta Q_{ijkl} = \mu + E_i + C_j + P_k + H_l + I_{ijkl} + \varepsilon_{ijkl} \tag{2.1}$$

which expresses the change in discharge (ΔQ_{ijkl}) as the mean change (μ) modulated by the main effects of four factors, the emission scenario (E_i, i = RCP2.6, A1B, A2), the climate model estimate (C_j, j = lower, medium, upper), the postprocessing (P_k, k = deterministic, probabilistic), the hydrological model (H_l, l = HBV, PREVAH, WaSiM), as well as the sum of the significant interactions between these factors (I_{ijkl}) and the residual error (ε_{ijkl}). The interaction terms allow accounting for nonadditive effects, i.e., for situations in which the combined effect of two factors is not the sum of their individual effects. We only considered first-order interactions, i.e., interactions between two factors, as accounting for and interpreting higher-order interactions is hard to physically justify. The assessment of the significance of first-order interactions, and their inclusion or not in the ANOVA model was based on F-tests. The sum of squares of each element (main effects, interactions, and error term) was divided by the total sum of squares of DQ to compute the fraction of variance explained by this element (von Storch and Zwiers, 2001; Bosshard et al., 2013a). This analysis was performed for each catchment, future period, and for both summer (JJA) and winter (DJF) projections.

2.2.7 Natural variability under present climate

To assess the significance of the projected temperature, precipitation, and discharge changes, we compared them with an estimate of the natural variability obtained by bootstrapping of observations. This estimate can be seen as noise (N) and was used to normalize the change of discharge (S), transforming it into a unit-less signal-to-noise ratio (S/N as, e.g., in Hawkins and Sutton, 2012). We followed Bosshard et al. (2011) to estimate N and constructed 100 thirty year time series by resampling years with replacement from the 1980-2009 records. This approach relies on the assumption that a baseline distribution can be estimated from these 30 years and that realizations equally likely under this climate can be created by resampling from this distribution. Then, 500 pairs of synthetized time series were randomly selected and the differences between these pairs computed. The standard deviation among these 500 differences was then used as an estimate of N. In the following, N is considered as the typical difference between two time series in absence of climate change, i.e., as a result of climate natural variability. Therefore, the standard deviation among the 500 differences of the annual, seasonal, and monthly means was computed, respectively. N was computed at the seasonal scale (winter and summer) for precipitation and temperature, and at the yearly, seasonal, and monthly scale for discharge. As periods of 30 years were considered, N reflects the natural variability over decadal time scales. The years used to construct the 100 time series and the subsequent 500 combinations were identical for the six catchments to account for intersite correlations. The Verzasca is the only exception, as its discharge record begins shortly before the beginning of hydrological year 1990, the thirty year time series were synthetized by bootstrapping measurements from the 1990-2009 period.

2.3 Results

2.3.1 Changes in the hydrological regimes and in glacier coverage

For each catchment, the agreement among the model chains of our setting is depicted in Figure 2.6 for 2070- 2099. The ensemble confirms climate change impacts on discharge already reported in previous studies on Swiss catchments (Horton et al., 2006; Bundesamt für Umwelt BAFU (Eds.), 2012; Köplin et al., 2012; Bosshard et al., 2013a) and in addition

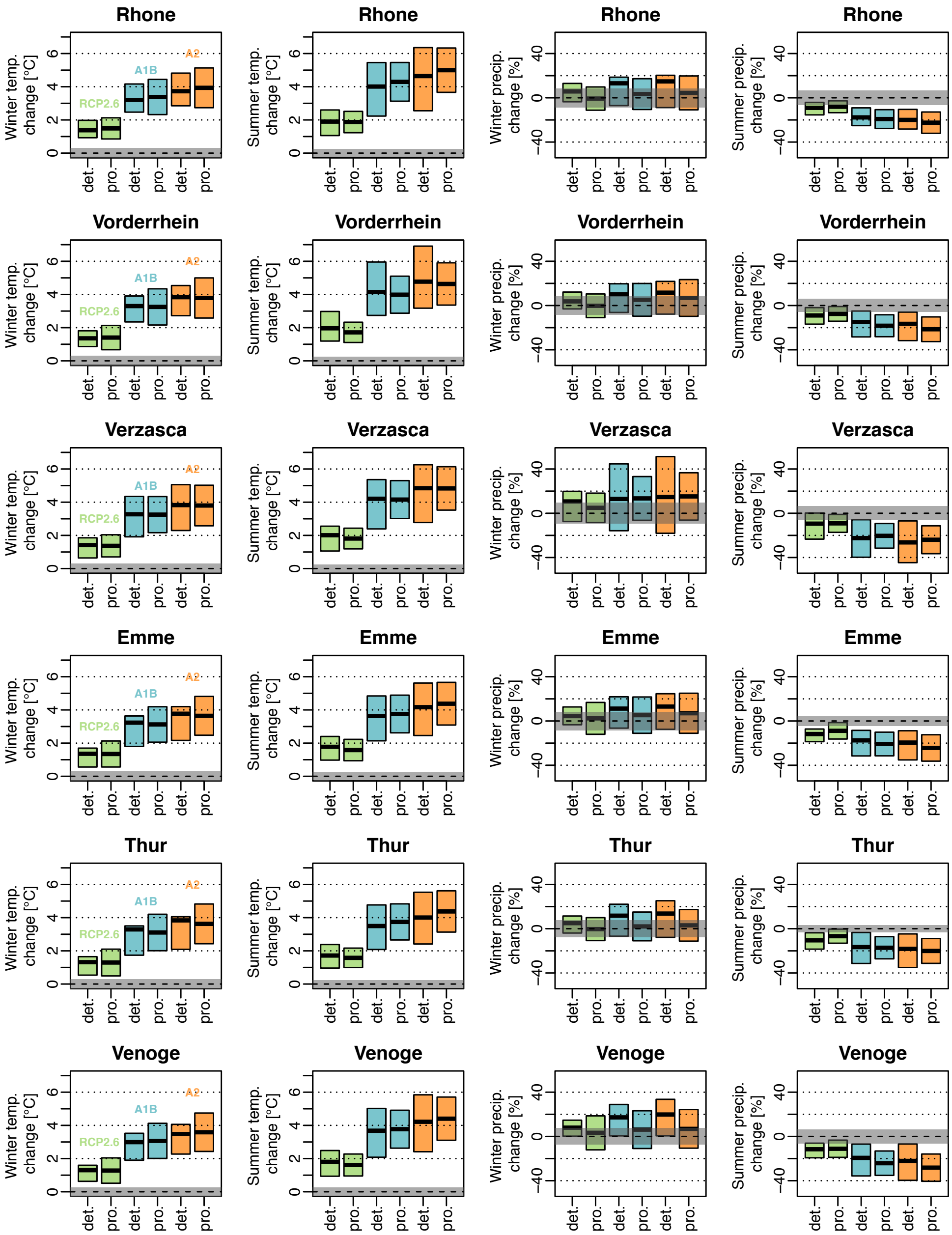

Figure 2.5: Projected changes in temperature and precipitation between 1980-2009 and 2070-2099 averaged over the catchments for winter (DJF) and summer (JJA). Estimates from the deterministic (det., Bosshard et al., 2011) and probabilistic (pro., Fischer et al., 2012) data sets are shown. For the deterministic data set, the lower and upper bounds of the boxes correspond to the minimum and maximum among the 10 GCM-RCM chains. For the probabilistic data sets, they correspond to the 2.5 and 97.5% quantiles. The thick black bars designate the median in both cases. The colors correspond to the emission scenarios RCP2.6 (green), A1B (blue), and A2 (orange). The gray shaded areas correspond the natural decadal variability estimated by bootstrapping precipitation and temperature records (±N).

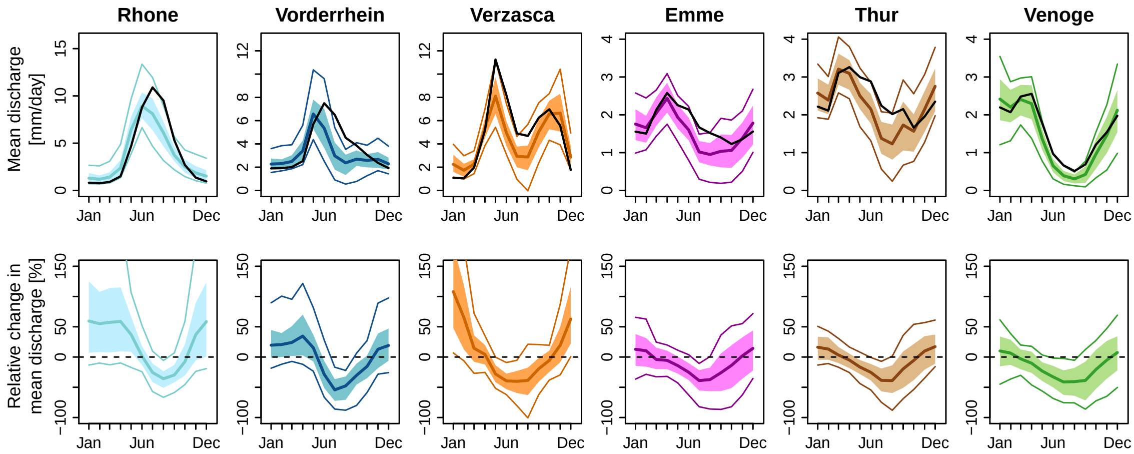

Figure 2.6: Projected regime changes for 2070-2099. The mean of the projections is represented by the thick colored line, the likely range (colored area) encompasses two thirds of all 54 model chains, and the minimum and maximum are shown by thin colored lines. The reference discharge (1980-2009) is depicted by a black line in the top row. The bottom row shows the relative difference between the reference and the projections. In Figures 6, 9, and 11, the catchments are ordered according to their mean elevation, from (left) the highest to (right) the lowest.

reveals strong model agreement for three of these impacts by the end of the century. First, for all catchments, lower summer (JJA) flows are projected by more than 90% of the model chains, irrespective of the emission scenario, climate model, postprocessing, or hydrological model. This is mainly related to a summer precipitation decrease (Köplin et al., 2012), as the changes in actual evapotranspiration were found to play a secondary role (Köplin et al., 2013; Adam et al., 2009). Second, more than 85% of the simulations indicate an earlier timing of spring-summer peak discharge, as a consequence of temperature increase. Third, larger winter (DJF) flows are projected by about two thirds of the runs in three lowest basins, and by more than 80% of them in the three highest basins. This mainly results from the higher fraction of liquid to solid precipitation during winter, leading to higher direct runoff. Note that the smaller snow storage causes lower melt peaks, especially at higher elevations.Winter precipitation might increase and contribute to higher winter discharge, but this remains uncertain, as the projected changes mostly fall within the estimated range of natural variability (Figure 2.5). Note that in contrast, the temperature changes simulated at the regional scale clearly emerge from noise in both winter and summer, as already reported by Bosshard et al. (2011) and Fischer et al. (2012). Although the expected changes in the Swiss plateau (Venoge) and prealpine catchments (Emme, Thur) are comparatively small in absolute terms, the relative changes are considerable (Figure 2.6, bottom row). In particular, the discharge during the low flow period, in summer, is likely to decrease by 25-45% for these three catchments.

The distributed modeling of glacier retreat forced by the probabilistic climate scenarios led to the projections summarized in Figure 2.7. By the end of the 21st century, and for the medium climate scenario, the glacierized area in the Rhone catchment is projected to be reduced to ~45% of its 2010 extent under RCP2.6, and to ~27% under A1B and A2. In the Vorderrhein catchment, the relative loss is higher, the glaciers are expected to retreat to ~17% and ~7% of their 2010 area, respectively, and only cover a few km2 by

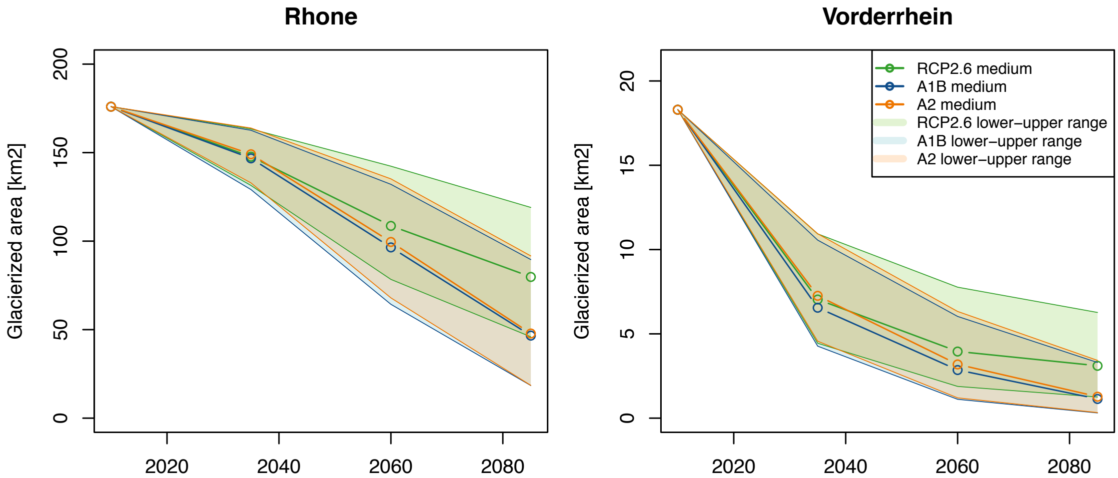

Figure 2.7: Projected evolution of glacierized area, for the lower, medium, and high estimates of the probabilistic climate projections under the RCP2.6, A1B, and A2 emission scenarios. Note the different scales of the y axes.

the end of the century. The uncertainty stemming from the GCM-RCMs is represented by the shaded areas and is large under each emission scenario. For instance, in the Rhone catchment, the lower and upper climate estimates (95% confidence interval) under RCP2.6 lead to a remaining glacierized area stretching from 26% to 68% of its 2010 value. Note that the uncertainties related to the formulation and the parameter values of the glacier model were not assessed in this paper.

2.3.2 Uncertainty decomposition

To quantify the respective contribution of the four uncertainty sources to the uncertainty captured by our experiment, analyses of variance (ANOVA) were performed, which allowed for uncertainty partitioning for each of the three future periods and for each catchment. We first assessed the significance of the main effects and interactions of the ANOVA model. For each main effect or first-order interaction, the null hypothesis is that its coefficients are all null, e.g., that the emission scenario has no influence on the projected discharge change ($E_i = 0$ for all i in Eq. 2.1). The hypothesis was evaluated using F-tests, and their p-values were summarized in Figure 2.8. We kept all the main effects and most of the interactions, including the interactions H:E and C:E because they are in most cases significant for the middle and end of century projections. The interactions P:E and H:P were excluded because they are not significant in most cases.

The uncertainty stemming from the GCM-RCMs, i.e., the combination of model uncertainty and natural climate variability, dominates in general (Figure 2.9). Our setup shows that the relative importance of this source varies with catchment characteristics. While the projection uncertainty is principally driven by the GCM-RCMs in the nonglacierized catchments, the hydrological models explain a comparable proportion of the variance in the partially galcierized catchments, Rhone and Vorderrhein. The different hydrological models lead to different projections in these catchments, in particular for the change in Rhone summer discharge around 2085, with HBV and WaSiM simulating a much larger (~2.8 mm/d) mean decrease than PREVAH (~0.5 mm/d). For the other catchments, the

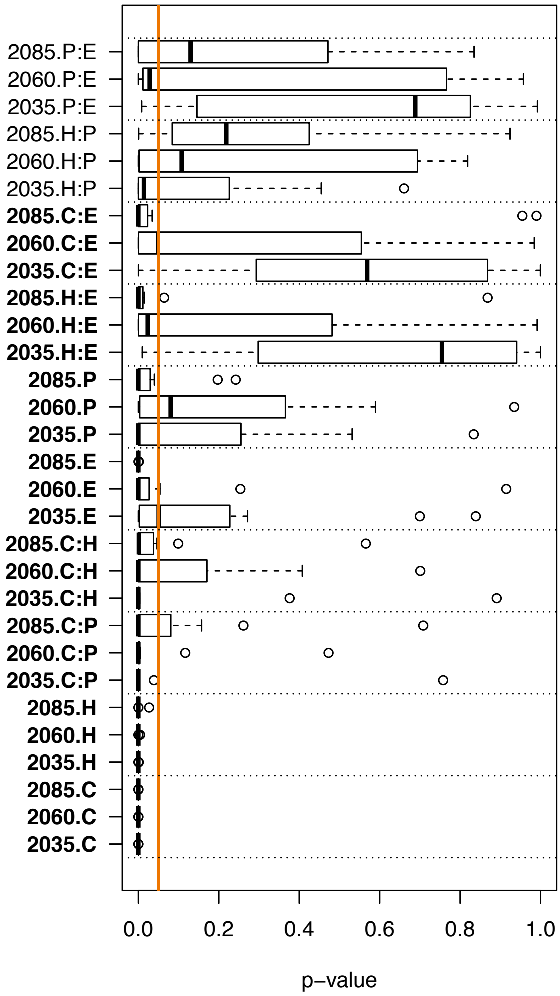

Figure 2.8: Significance of the ANOVA model elements. The boxplots summarize the p-value of the F-tests for the main effects and first-order interactions for changes in winter (DJF) and summer (JJA) in the six catchments, i.e., a total of 12 values per boxplot. The four factors are the emission scenario denoted by E, the climate model C, the postprocessing P, and the hydrological model H (see Eq. 2.1). First-order interactions between factor X and factor Y are denoted by X:Y. The orange line indicates p-values of 5%. The factors and interactions considered in the final ANOVA model have their names in bold.

hydrological models explain little of the variance, which means that, in these catchments, the differences in model structure, resolution, and calibration barely influence the discharge projections.

The fraction of variance explained by the emission scenarios increases with time, for both seasons and in all catchments, and so does the significance of the emission scenario factor in the ANOVA model (Figure 2.8). This source of uncertainty explains a higher fraction of variance in summer than in winter, but is usually not the dominant one by the end of the century for the catchments investigated. The difference in complexity between the two postprocessing methods, and the different sets of climate models they rely on, appear to have little influence. Finally, the variance explained by the interactions needs to be considered, in particular in summer.

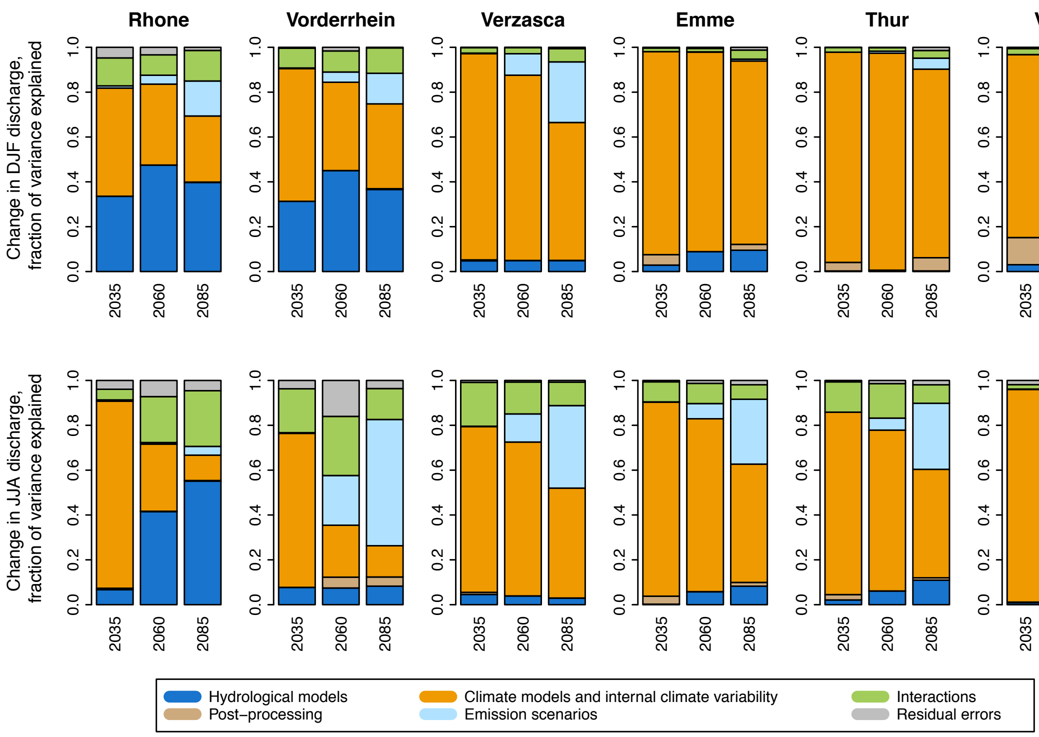

Figure 2.9: Decomposition of the projection variance. ANOVA partitioning among the four sources of uncertainty, the significant interactions, and the residual errors. Results for discharge changes in (top) winter (DJF) and (bottom) summer (JJA) are shown for the three 30 year future periods, centered on 2035, 2060, and 2085.

2.3.3 Emergence of the climate change signal from natural discharge variability

The annual discharge cycles of the time series constructed by bootstrapping are shown in Figure 2.10. Overall, the variability is higher in lower-elevation basins, as illustrated by the higher dispersion of the 100 constructed time series around the cycle directly derived from observations. When computed for each month, the variability corresponds in average to 7-19% of the discharge on the same month, with lower values typically reached in higher-elevation catchments. A clear exception to this relation is the Verzasca, which exhibits a rather high variability given its elevation. This relationship between elevation and variability translates to the HFD, which presents higher variability at lower elevation. Its standard deviation ranges from 1.8 days in the Rhone catchment to 8.1 days for the Venoge, the Verzasca being again an outlier with a value of 11.8 days. As noted earlier, higher- elevation catchments are associated to a later HFD.

Estimating the natural discharge variability in the six catchments allows for the investigation of the emergence of the climate change signal. It already emerges from noise in certain months of the first future period (2020-2049) in the alpine catchments (Rhone and Vorderrhein) and irrespectively of the emission (Figure 2.11). The impacts become more severe with time and our results indicate that at the end of the century, the climate change signal clearly emerges from natural variability for all catchments in summer, even if emissions are limited to the lowest level (RCP2.6). The S/N ratios are higher in the higher-elevation catchments.

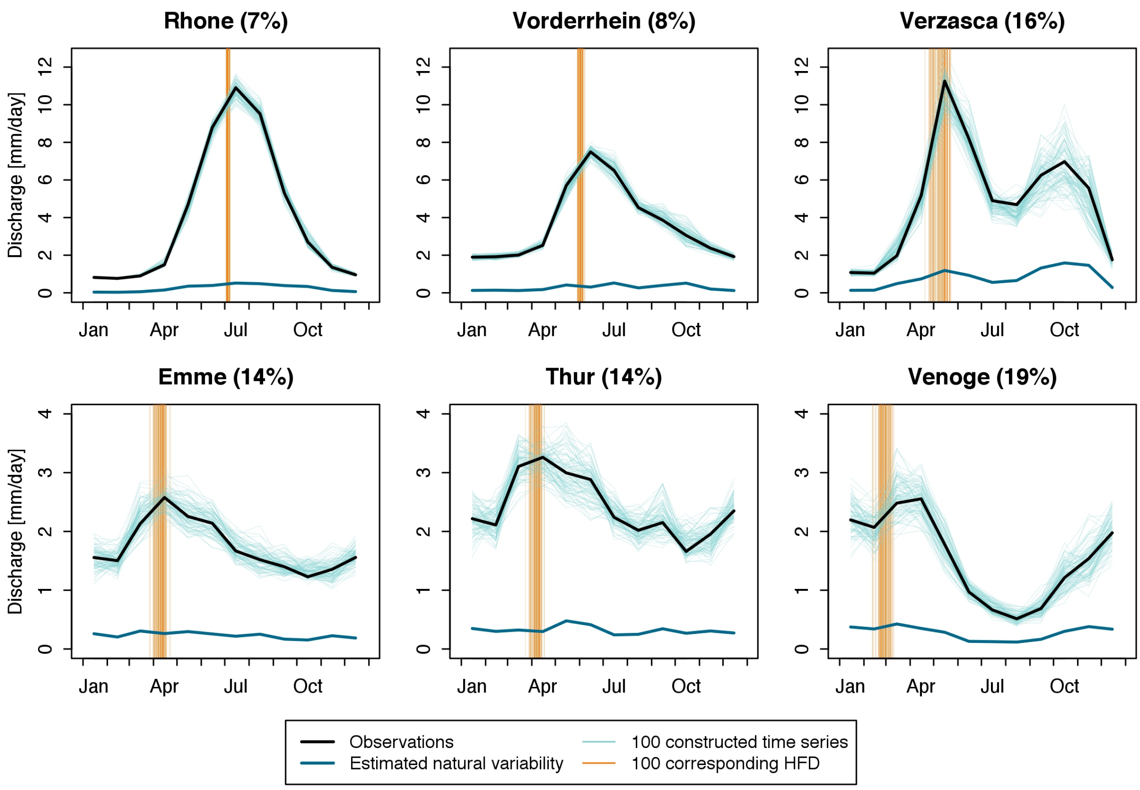

Figure 2.10: Natural discharge variability estimated by bootstrapping. Annual discharge cycles obtained from the observations and from the 100 time series constructed by bootstrapping. The natural variability N, estimated for each month from the 100 realizations, is depicted by the thick blue line and its mean value relative to the observed discharge is indicated in brackets after the catchment name. The HFD derived from the 100 realizations are indicated by vertical orange lines.

2.3.4 Sensitivity of the impacts on discharge to the emission scenarios

We explored the influence of the emission scenarios on the hydrological regime by computing the average discharge change for each emission scenario, catchment, and future period. As expected from the analyses of variance, the differences between the emission scenarios increase with time (Figure 2.11). These differences are negligible for the first future period (2020-2049), the impacts occurring independently of the emission policy or lack thereof. Differences between the intervention (RCP2.6) and the nonintervention (A1B and A2) scenarios become clear for the 2045-2074 projections, when impacts increase (not shown). By the end of the century (2070-2099), a further increase of the climate change impacts is projected, but there is only little difference between A1B and A2, which is consistent with the similarity of their temperature and precipitation projections for the study basins (Figure 2.5). An important result is that, according to our simulations, following the intervention scenario would reduce the largest impacts on the regime (in summer and winter) by the end of the 21st century by about a factor of two.

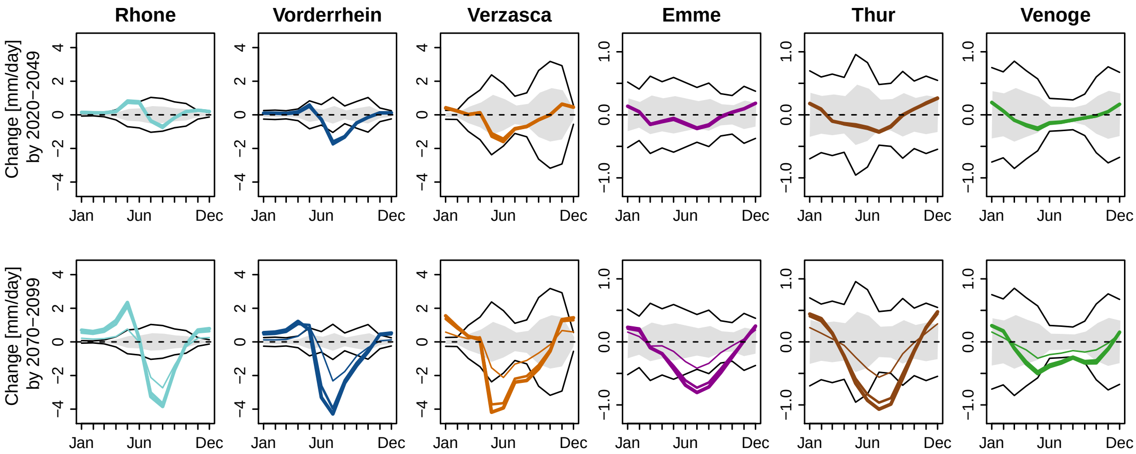

Figure 2.11: Changes per emission scenario compared to the natural variability. The two thick colored lines represent the projected changes under the two nonintervention scenarios (A1B and A2), the thin colored line the changes under RCP2.6. Each of them is the average of 18 discharge simulations. The gray shaded areas and black lines depict the natural discharge variability estimated under present climate ($\pm N$ and $\pm 2N$, respectively). The climate change signal is considered to emerge ($S/N > 1$) when the colored lines leave the gray shaded area.

2.4 Discussion

2.4.1 A large but inevitably incomplete set of model chains

This study relies on a large set of model chains, with 3 emission scenarios, 10-20 GCM-RCMs, 2 postprocessing methods, and 3 hydrological models combined factorially and applied to six catchments. We nevertheless acknowledge that model sampling is neither complete nor random, which is a general barrier to a full uncertainty sampling (Knutti et al., 2010). The models are also equally weighted, in a finite number and somewhat interdependent (see the discussion on climate models in Masson and Knutti, 2011, and the hydrological model characteristics in Table 2.2). It thus cannot be excluded that some agreement between the simulations stems from model similarities, such as a common bias among the climate or hydrological models, and hence should not be interpreted as low uncertainty. It is debatable whether adding more chains would change our characterization of the projection uncertainty, improve our understanding of future catchment discharge, and provide a stronger support for decision making under uncertainty. Here we argue that even though sets of model chains are inevitably incomplete, the use of a wide and balanced set of combinations provides new insights into the influence of the catchment on the uncertainty of the discharge projections, relates to fundamental questions on the relation between model complexity and performance, and helps to find robust features that can better support decision making. These points are discussed below.

2.4.2 How much do catchments determine uncertainty partitioning?

The decomposition of variance revealed that the discharge uncertainty is in general dominated by the uncertainty from the GCM-RCM projections, which is in agreement with previous studies (Wilby and Harris, 2006; Prudhomme and Davies, 2009; Dobler et al., 2012; Bosshard et al., 2013a; Huss et al., 2013). Our setup further allows for the investigation of changes in the uncertainty partitioning from catchment to catchment (Figure 2.9).

The hydrological models explain a considerable portion of the variance in the Rhone and Vorderrhein catchments, but their influence is much smaller in the other catchments. In our view, this results from key characteristics of these two catchments that differentiate them from the others: their nival to glacial regime, their more complex topography, and the presence of dams. In alpine catchments, a realistic regime representation is conditioned by the realistic simulation of accumulation and melt processes of snow, and of ice in glacierized catchments. These processes are subject to structural uncertainties (they are formulated in different ways in the hydrological models, see Table 2.2) and parametric uncertainties (parameter estimation is performed using different methods, see Table 2.2, and is challenging as a result of equifinality (Beven, 2006) and of the highly multidimensional parameter space (Kirchner, 2006)). Overall, the value of parameters regulating snow and ice is poorly constrained when model calibration is performed on the basis of simulated discharge alone. This can lead to compensation errors, which then contribute to differences in the discharge simulated by the different models. Clearly, structural and parametric uncertainties also influence discharge simulations in lower-elevation catchments, but the potential for compensation errors resulting in misallocations of water between glacier, snow, and runoff is arguably larger at higher elevation, where storage of water as ice and snow is higher. A better agreement between the models might be achieved by multicriteria calibration, involving, for instance, the incorporation of glacier mass balance and snow cover extent retrieved from satellite imagery, as implemented by Parajka and Blöschl (2008) and Finger et al. (2012). This would enable snow and ice mass balances to be more realistically simulated, an aspect not evaluated by calibration methods based on discharge simulations alone. Another characteristic influencing uncertainty partitioning is topography. Although the hydrological models are forced by the same climate projections, the more complex topography in higher-elevation catchments leads to higher errors in the interpolation of these projections, an operation performed by the hydrological models. Finally, the discharge perturbations induced by hydropower production during the calibration period decrease the identifiability of the hydrological model parameters, as the operational rules were not implemented in this study. These numerous factors contribute to explain why the uncertainty stemming from the hydrological models is higher in the alpine catchments.

Catchments also influence the importance of the emission scenario. In winter, it is lower in the Thur, Emme, and Venoge basins (mainly precipitation driven) than in the higher-elevation basins, Rhone, Vorderrhein, and Verzasca (more temperature driven). This is consistent with the climate projections for winter, which are similar under the three emission scenarios for precipitation, but differ for temperature (Figure 2.5). The difference between low and high-elevation catchments is reduced in summer, when both precipitation and temperature changes depend strongly on the emission scenarios.

The influence of the postprocessing method and the set of GCM-RCM chains they rely upon does not seem to depend on catchment characteristics and is weak in all cases. This reflects the great similarity of the projections from the deterministic and probabilistic data sets for the study basins (Figure 2.5). The differences in the projected discharges could however be stronger for extremes (Wilby and Harris, 2006; Dobler et al., 2012) and if other postprocessing methods, such as quantile mapping (Dosio and Paruolo, 2011), were also considered. Also, if a smaller number of GCM-RCMs had been used, the selection of the climate models might have had a larger effect on the discharge projections. Wilby and Harris (2006), for instance, used four GCMs and showed that, while three of them lead to rather similar changes in future low flows, the last one resulted in quite different projections (their Figure 5b). In such a case, including this last GCM in the analysis or not, can lead to quite different conclusions about the origin of the uncertainty in the discharge projections.

It should be stressed that the validity of the uncertainty partitioning relies in particular on the experimental design. A critical point here is the assessment of the interactions terms. The two most significant interactions are between the climate model, the postprocessing method (C:P), and the hydrological model (C:H, Figure 2.8). The combined effect of these elements on discharge is thus nonadditive. For instance, when the medium climate estimate and the hydrological model HBV are combined, their effect on the discharge projection is not only the sum of their respective effects, but rather this sum corrected by an additive interaction term (Eq. 2.1). An important consequence is that, if the uncertainty stemming from the chain elements (e.g., from the hydrological models) had been computed by varying the models for this element (e.g., using sequentially HBV, PREVAH, and WaSiM) while keeping the rest of the chain constant (e.g., using only SRES A1B, medium climate estimate, and the probabilistic postprocessing), a biased estimate of the uncertainty could result, as significant interactions (e.g., between the climate and the hydrological models) would not have been sampled. The nonnegligible role played by the interactions hence stresses that the choice of the setup is key for a reliable assessment of uncertainties in the modeling chain. Note that although we rely here on a factorial design, a less computing intensive approach based on a fractional factorial design requiring only a subset of the runs could be envisaged for future studies (Wu and Hamada, 2009).

2.4.3 Which model structure and calibration for hydrological impact studies?

In the nonglacierized catchments, the proportion of the variance explained by the hydrological models is almost negligible (Figure 2.9). The three models (HBV, PREVAH, WaSiM) rely on different spatial discretizations (100 m elevation bands, hydrological response units or 250 x 250 m^2 grid), different parameter estimation procedures (genetic algorithm, regionalization, PEST), different numbers of land cover classes (1, 22 and 25), and different reservoir structures. Yet their projected regimes seem barely sensitive to these differences. We see three complementary explanations of this outcome.

First, external forcing, i.e., atmospheric projections, can act as the main driver of the simulated discharge, overwhelming any differences between the hydrological models. Climate projections showing significant differences from the climate during the reference period (Figure 2.5), our results suggest that these differences can overrule the structural differences among the hydrological models.

Second, our study focused on changes in the mean seasonal discharge over 30 year periods. A model might perform better at capturing individual events, for instance, the few large flooding events during this period, but this is averaged out when considering changes of mean discharge. Larger differences between the models might thus have emerged if other components of the hydrograph had been systematically investigated, for example, low flows (Velázquez et al., 2013).

Third, the similarity of the projections delivered by different models raises the question of how much complexity is needed in hydrological models executed under climate change. This is related to the much-debated relation between model performance and structure, which can be here formulated as: which hydrological model elements are necessary to deliver reliable discharge simulations in a changing climate? The answer arguably depends on the catchment characteristics and on the hydrological parameter investigated, as there is most likely no model structure that provides reliable projections in all cases (Savenije, 2009). Several contributions exist to support the choice of model structure. They rely, for instance, on model intercomparisons (Smith et al., 2004, 2012), on the Framework for Understanding Structural Errors (FUSE, Clark et al., 2008; Staudinger et al., 2011), on differential split sample tests (Seiller et al., 2012), on a model evaluation without calibration method (Vogel and Sankarasubramanian, 2003), or adopt a more process-oriented approach (Herman et al., 2013). Concerning hydrological model calibration, Merz et al. (2011) and Coron et al. (2012) highlighted the issue of parameter stationarity over time and discuss the consequence thereof for simulation of hydrological impacts. Others consider the whole impact modeling chain (Figure 2.1) to estimate the sensitivity of hydrological projections to the model structure (Velázquez et al., 2013) or to the method chosen to estimate parameters value (Poulin et al., 2011). These later studies consider the hydrological model in the context of the cascade of uncertainty and compare its contribution to that of the other elements of the model chain, hence providing particularly valuable insights for the further development of hydrological modeling in changing conditions.

2.4.4 Why is natural discharge variability higher in lower-elevation catchments?

The overall negative correlation between catchment elevation and natural discharge variability under present conditions (Figure 2.10) can be explained by a damping by glaciers and the snowpack. According to Lang (1986), this damping can occur through three main mechanisms. During dry years characterized by lower than normal precipitation and higher net radiation and sensible heat, glacier balance is usually negative and the contribution of meltwater to discharge is significant. Conversely, during years with much precipitation, the input to glaciers is usually higher and melt is lower, which corresponds to a positive glacier mass balance and a smaller contribution to discharge. Glacier melt hence tends to compensate for year-to-year rainfall anomalies. Another compensating effect stems from the tendency of climatic situations leading to higher glacial melt to also cause higher evapotranspiration in basins spanning over a wide range of elevations. Finally, the snowpack can also damp discharge variations, by temporarily containing intense precipitation events and then releasing the water progressively.

The comparatively high discharge variability of the Verzasca given its elevation probably comes from the location of the basin, on the southern flank of the Alps. Its climate is directly influenced by Mediterranean airflows and the natural precipitation variability of its region is larger than on the northern side of the Alps (Fischer et al., 2012). Further, when comparing the projected changes to the natural variability (Figure 2.11), it appears that the changes in regime are more pronounced in the higher-elevation catchments, a result also reported by Fatichi et al. (2014).

2.4.5 Can the identification of robust changes help adaptation?

We investigated the robustness of the projections by considering robustness as the combination of a strong agreement between simulations and a projected change exceeding natural variability (Knutti and Sedláček, 2012). These two aspects were investigated separately in sections 2.3.1 and 2.3.3, respectively. To conclude this study, they were combined in Figure 2.12 for the three regime impacts discussed above: lower summer discharge, higher winter discharge, and change toward more rain-dominated regime. In addition, we also considered changes in the mean annual volume, which correspond to changes in the water balance integrated over 30 years. In the four cases, the robustness of the projections is higher for the higher-elevation catchments, i.e., their mean future change normalized by the natural variability is higher, and more model chains agree on the sign of the change. We explain this outcome by the robustness of the future climate simulations, which is higher for temperature than for precipitation (Figure 2.5), resulting into more robust discharge changes in temperature-driven, higher-elevation catchments. Similarly, as the robustness of the climate change signal tends to increase with time in the catchments investigated (Figure 2.5), so does the robustness of the hydrological projections.

Note however that different robustness patterns emerge. In the lower-elevation catchments, there is a strong model agreement and high signal-to-noise ratios about lower summer discharge, but the projections confidence is weaker when it comes to higher winter discharge, as could be expected from the results shown in Figure 2.6. Further, note that model agreement on the shift toward more rain-dominated hydrological regimes is particularly high and is already complete (100%) in 2035 for the three highest catchments. Overall, high signal-to-noise ratios are reached for these three first changes. It implies that the projected changes are unlikely to stem from natural variability alone, but are rather a consequence of anthropogenic emissions.

According to our ensemble of simulations, the total discharge volume is expected to decrease in the future. This change is less robust than the other three (Figures 2.12a-2.12c versus Figure 2.12d), yet by 2085 it emerges from natural discharge variability and is projected by 60-85% of the model chains. Possible causes include the expected decrease in summer precipitation and smaller glacier contribution to discharge, and perhaps the precipitation shift from snow toward rain (Berghuijs et al., 2014). At this stage, the respective contribution of these processes is unclear.

Because of the diversity of the model combinations considered, our projections are subject to substantial uncertainties, as shown by the large spread of the ensemble (Figure 2.6). Nevertheless, Figure 2.12 shows that by the end of the century, robust changes in regime emerge from the noise, that they are supported by a large agreement among the model runs and that they are valid across the whole range of catchments. We argue that identifying such features can contribute to support decision making on adaptation strategies.

2.4.6 Can our results and method be generalized?

Our study focused on six catchments, but data sets with hundreds of catchments are becoming more easily available (Gupta et al., 2014). They can be combined with climate projections to performa similar systematic analysis in a larger number of catchments spanning a wider range of climatic conditions. However, since uncertainty sampling and decomposition requires a large number of simulations, available computing infrastructures will often limit the analysis to a selection of catchments. So which catchments to select? A solution is to reduce the dimensionality of the problem. Köplin et al. (2012), for instance, run a single hydrological model under a single emission scenario using 10 GCM-RCMs postprocessed by a single method in 186 mesoscale Swiss catchments. A cluster analysis then enabled them to reduce these 186 catchments to 7 response types, from which the catchments used in this study were chosen (Figure 2.2). We then produced hydrological projections for this selection of catchments using several emission scenarios, hydrological models and GCM-RCM postprocessing methods. The similarity of our results for the Emme, Thur, and Venoge catchments is in agreement with their clustering, as they associated these basins to the same type of response to climate change. This suggests that our results may be extrapolated to other Swiss watersheds on the basis of their cluster analysis

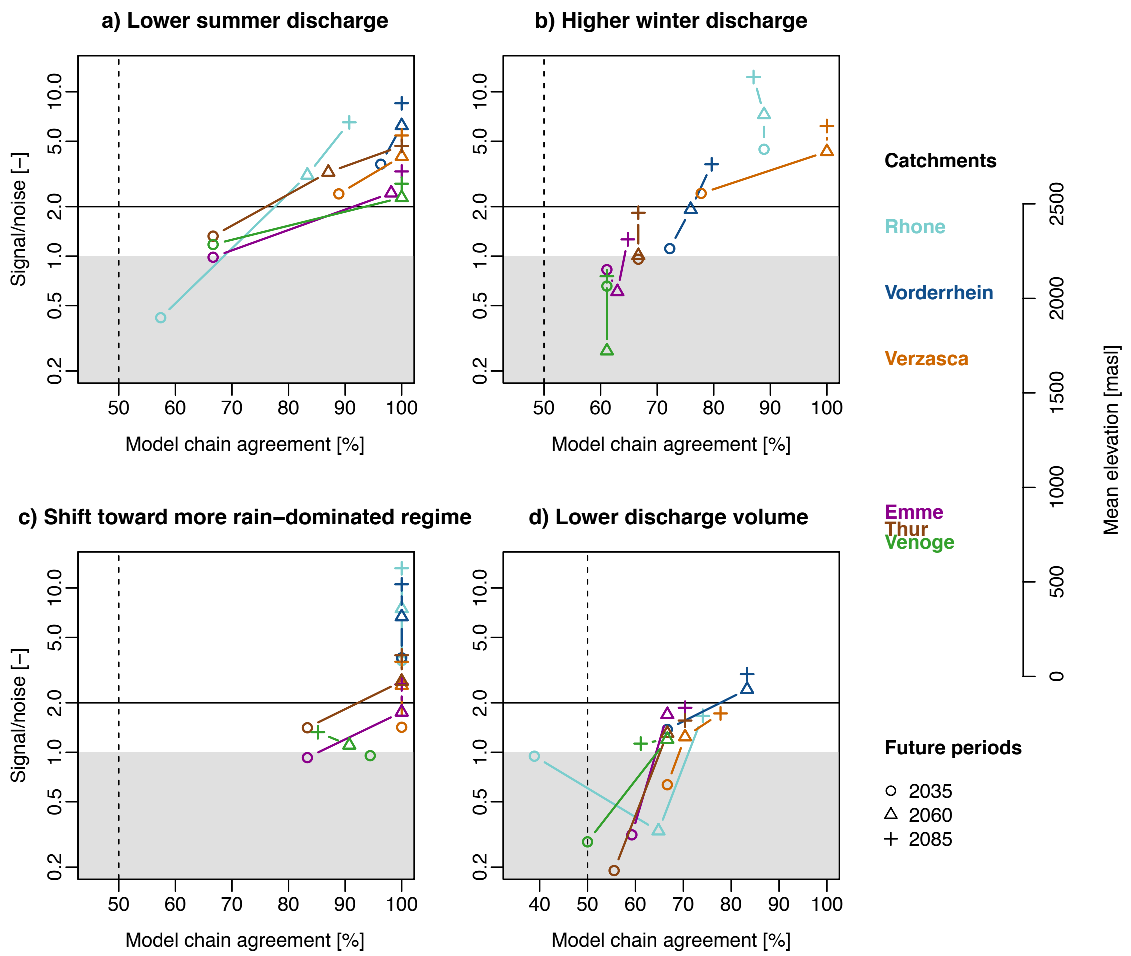

Figure 2.12: Robustness of the four regime changes. Agreement on the sign of the change among the 54 model chains (x axis) and mean change computed from the 54 chains normalized by the estimated natural variability (significance, y axis). The symbols refer to the future periods and the colors refer to the catchments (see legend on the right, where the catchments mean elevation is also indicated). The symbols located closest to the upper right corner of the plots denote the most robust projections. The gray area and the horizontal black line indicate S/N values below 1 and 2, respectively. The lowest possible degree of agreement (50%) is indicated by the vertical dashed line. A shift toward more rain-dominated regime is identified by an earlier half-flow date. Note the logarithmic scale of the y axis and the wider interval on x axis for Figure 2.12d.

and advocates for the application of a similar approach to other regions.

2.5 Conclusions

This study considers and systematically analyses a large number of uncertainty sources when simulating future hydrological regimes in Swiss catchments. Analyses of variance based on a factorial experimental design were used to decompose the uncertainty of the projections. This showed that although GCM-RCMs are usually the main source of uncertainty, the uncertainty stemming from the hydrological models in the catchments dominated by snow and ice melt is substantial. In contrast, the choice of the hydrological model is barely significant in the lower-elevation basins. The importance of the emission scenario increased with time into the future, yet without becoming the major source of uncertainty. The postprocessing of the climate projections plays the least relevant role in our study.

We assessed the robustness of expected changes in regime already reported in the literature: the decrease of summer discharge, the increase of winter discharge and shift toward lower-elevation, more rain-driven regimes. These changes are characterized by a large agreement among the simulations and by projected changes significantly larger than the natural variability, which was estimated by the bootstrapping of discharge records. Given the wide range of models involved in our setup, we concluded that these changes are robust despite the uncertainty of the projections, especially in the higher-elevation basins.

We compared the projections under the intervention scenario RCP2.6, implying considerable efforts to reduce emissions, and under two nonintervention scenarios, SRES A1B and A2. Over the coming decades, the impacts are projected independently of the scenario and the climate change signal already emerges from natural discharge variability in the high-elevation basins. By the end of the century, the projected impacts are more pronounced, the climate change signal emerges in all basins in summer, and there are clear differences between the intervention and the nonintervention scenarios. Our results indicate with confidence that impacts on discharge can be reduced considerably if a stringent emission policy is adopted.

This study relies on a coordinated modeling experiment, enabling the investigation of projection uncertainties in a systematic and quantitative way, and in several catchments. This hydrological modeling framework provides new insights into future hydrological regimes, and because it allows for the identification of robust changes, we argue that it is an important step for the support of decision making based on uncertain projections.

Acknowledgements

We thank the CH2014 Impacts project (CH2014 Impacts, 2014) for providing the framework in which this study was initiated. We are indebted to the Center for Climate Systems Modeling (C2SM) for producing and providing the CH2011 data. This data set relies on the climate simulations from the EU-funded FP6 Integrated Project ENSEMBLES whose support is gratefully acknowledged. We thank Reinhard Furrer for his help with the statistical analysis, Thomas Bosshard and Luzi Bernhard for inspiring discussions, Jan Schwanbeck for his help with the creation of PREVAH input files, as well as the Associate Editor and the three reviewers for their comments, which contributed to improve this paper. We acknowledge the Swiss Federal Offices for the Environment (FOEN), of Meteorology and Climatology (MeteoSwiss) and of Topography (swisstopo) for the hydrological, meteorological, and topographic data. This study was partially funded by the Swiss National Science

Foundation (grant 200021_131995). The discharge projections produced in this study are available upon request from the first author.

Chapter 3

The influence of natural variability and interpolation errors on bias characterization in RCM simulations

Nans Addor[1], Erich M. Fischer[2]

[1] Department of Geography, University of Zurich, Zurich, Switzerland
[2] Institute for Atmospheric and Climate Science, ETH Zurich, Zurich, Switzerland

Published in Journal of Geophysical Research: Atmospheres, 2015
doi:10.1002/2014JD022824

Abstract

Climate model simulations are routinely compared to observational data sets for evaluation purposes. The resulting differences can be large and induce artifacts if propagated through impact models. They are usually termed 'model biases', suggesting that they exclusively stem from systematic models errors. Here we explore for Switzerland the contribution of two other components of this mismatch, which are usually overlooked: interpolation errors and natural variability. Precipitation and temperature simulations from the RCM COSMO-CLM were compared to two observational data sets, for which estimates of interpolation errors were derived. Natural variability on the multi-decadal time scale was estimated using three approaches relying on homogenized time series, multiple runs of the same climate model and bootstrapping of 30-year meteorological records. We find that although these methods yield different estimates, the contribution of the natural variability to RCM-observation differences in 30-year means is usually small. In contrast, uncertainties in observational data sets induced by interpolation errors can explain a substantial proportion of the mismatch of 30-year means. In those cases, we argue that the model biases can hardly be distinguished from interpolation errors, making the characterization and reduction of model biases particularly delicate. In other regions, RCM biases clearly exceed the estimated contribution of natural variability and interpolation errors, enabling bias characterization and robust model evaluation. Overall, we argue that bias correction

of climate simulations needs to account for observational uncertainties and natural variability. We particularly stress the need for reliable error estimates to accompany observational datasets.

3.1 Introduction

General circulation models (GCMs) and regional climate models (RCMs) are traditionally evaluated by comparing their outputs to observational data sets. One common outcome of such exercise is that mismatches between model data and observations are reported as biases (Christensen et al., 2008; Dosio and Paruolo, 2011; Themeßl et al., 2011a; Boberg and Christensen, 2012). These biases can be large and when climate model output is directly used to drive impact models, such as hydrological models, it usually leads to biased simulations. This paved the way for the development of post-processing methods to enable the use of RCM simulations for impacts modeling (Fowler et al., 2007; Maraun et al., 2010; Hagemann et al., 2011; Rojas et al., 2011; Chen et al., 2013). Although the understanding of the reasons behind these biases is constantly progressing, which is contributing to the overall improvement of climate models with each new generation (Knutti et al., 2013), significant biases remain in state-of-the-art simulations, both at the global and regional scale (Mueller and Seneviratne, 2014; Kotlarski et al., 2014). It is hence expected that one will continue to resort to bias correction methods in the near future.

In parallel, recent studies showed that the importance of natural climate variability had not received enough attention when interpreting climate projections (Deser et al., 2012a,b; Fischer et al., 2013). Deser et al. (2012a) performed an ensemble of simulations with a single GCM started from slightly different atmospheric initial conditions. The spread across the different simulations provides a quantitative estimate of the natural variability of the climate system. The role of natural variability is particularly relevant at the regional to local scale, as well as for variables such as precipitation or any type of climatic extremes. In some regions of North America, they revealed considerable differences among the ensemble members, as a sole consequence of natural variability, preventing to determine even the sign of the precipitation change in those locations in the coming decades. A central characteristic of the uncertainty induced by natural variability is that it is ultimately irreducible, as it is unpredictable (i.e., aleatory) on time scales longer than a couple of decades. At least two questions emerge in the light of these findings. How much does natural variability influence the results of climate model evaluation over multi-decadal periods? Does the unpredictable nature of internal variability preclude our chances to reduce climate model biases by post-processing?

Analogously, it is at present unclear how errors in observational gridded data sets, used as reference for model evaluation and bias correction, compromise our efforts to reduce biases and to perform reliable model evaluation. Sunyer et al. (2013) underscored the importance of the reference data set when characterizing RCM performance and biases. They in particular show that biases amplitude, and sometimes even their sign, can depend on the observational data set. Similarly, Gómez-Navarro et al. (2012) demonstrated the ranking of climate models can vary if different observational references are used. Both studies stress the importance of accounting for observational uncertainties. Although several bias correction studies mention deficiencies of gridded data sets (e.g., Kyselý and Plavcová, 2011; Themeßl et al., 2011a) and methods exists to account for observation uncertainties in hydro-climatic studies (e.g., Bellprat et al., 2012; Greve et al., 2014), studies evaluating the quality of the observational data set before using it to bias-correct RCM simulations are rare. It is generally implicitly assumed that gridded data sets represent 'true' atmospheric conditions, or at least that their errors are small enough to be neglected. Although important efforts already went into the quantification of errors in observational data sets (Haylock et al., 2008; Hofstra et al., 2009; Morice et al., 2012), it is unclear how detrimental these errors are for the evaluation and bias correction of RCM simulations. In other words, how well can biases be characterized, and hence corrected, given the uncertainties in reference data sets?

To address these questions, we structured our study in two parts. First, we quantify multi-decadal natural variability, interpolation errors and RCM-observation differences using different techniques and data sets, then discuss the limitations and advantages of these approaches and compare their results. Second, we combine the results from the first part to explore the contribution of multi-decadal natural variability and interpolation errors to RCM-observation differences and discuss the implications for the evaluation and bias correction of RCM simulations.

3.2 Data and methods

3.2.1 The contribution of natural variability, interpolation errors and model biases

Here we consider that the difference D between RCM simulations and observations is the combined result of three sources of uncertainty under present climate: the bias of the climate model M, the error of the interpolation procedure I and the natural climate variability N. To explore the respective contribution of M, N and I to D, we used a simple method. We produced quantitative estimates of D, N and I and normalized D by N and I to assess how likely the RCM-observation mismatch is to stem from natural variability and interpolation errors, respectively. Our main assumptions can be summarized as follows. When D is within the uncertainty range associated to I and N, we consider that it is not necessarily caused by a model bias, but could originate from interpolation errors or from the natural variability, respectively. In contrast, we interpret high values of the $|D/N|$ and $|D/I|$ ratios as situations in which D emerges from N and I, meaning that it is too large to be explained by N and I alone. In those cases we assume that the mismatch results principally from its third component, M. Further, we consider that the larger the contribution of N to D, the more compromised model evaluation and bias correction are, as a consequence of the stochastic nature of N. We develop the consequences of high and low $|D/N|$ and $|D/I|$ in more detail when discussing our results in Section 4.

3.2.2 Spatial and temporal scales

RCM evaluation can be performed at different spatial scales. At regional scale, RCM grid cells are combined into regions, e.g., Scandinavia or the Alps, and compared with observation aggregated over the same areas (e.g., Christensen and Christensen, 2007; Bellprat et al., 2012). At grid scale, RCM simulations at each grid point are compared with observations interpolated on the same grid (e.g., Dosio and Paruolo, 2011; Themeßl et al., 2011a). At station scale, stations observations are compared with RCM interpolated at station location (e.g., Gudmundsson et al., 2012). Evaluations at regional scale allow to condense results and summarize model performance over large areas, while evaluations at grid and station scales provide more detailed insights into the models ability to capture smaller scale phenomena (e.g., the influence of the Alpine foothills on precipitation). These two smaller scales are especially relevant for impact studies, such as those relying on hydrological modeling at catchment scale (e.g., Dobler et al., 2012; Addor et al., 2014) and are the focus of this study.

The area of interest here is Switzerland because of the availability of a wide range of climatic data and areas of different topographic complexity. We used in particular temperature and precipitation records from a high-density station network, homogenized time series going back to 1864 and a set of RCM simulations designed to sample natural variability. Further, Switzerland features a low mountain range, the Jura, a plateau and a large Alpine area (Figure 3.1), which makes it an ideal test site to evaluate how well atmospheric processes are captured in gentle to complex topography situations. Our focus is on mean temperature and mean precipitation in winter (December, January and February, DJF) and summer (June, July and August, JJA). All results in continuation are for the reference period 1980-2009, unless indicated otherwise, and rely on daily values.

It is important to stress that a key reason behind the mismatch of station observations and gridded RCM simulations is the difference in spatial scale. To explore the influence of scale, we conducted our analysis at the station scale and at the grid scale, and compared the results.

3.2.3 Observational data sets

We used two sets of station measurements, as well as one gridded observational data set. We used measurements over 1980-2009 coming from 392 rain gauges and 86 temperature stations operated by the Swiss Federal Office of Meteorology and Climatology, MeteoSwiss. Each station had at least 80% of available daily measurements for the study period. In addition, we used measurements going back to 1864 from 11 temperature and 9 precipitation stations and homogenized to correct inconsistencies arising from changes in stations location and measuring practices (Begert et al., 2005). Further, we relied on recently released data sets of precipitation (RhiresD) and temperature (TabsD) interpolated by MeteoSwiss on a 2 km grid covering Switzerland. Those data products are analyses produced using

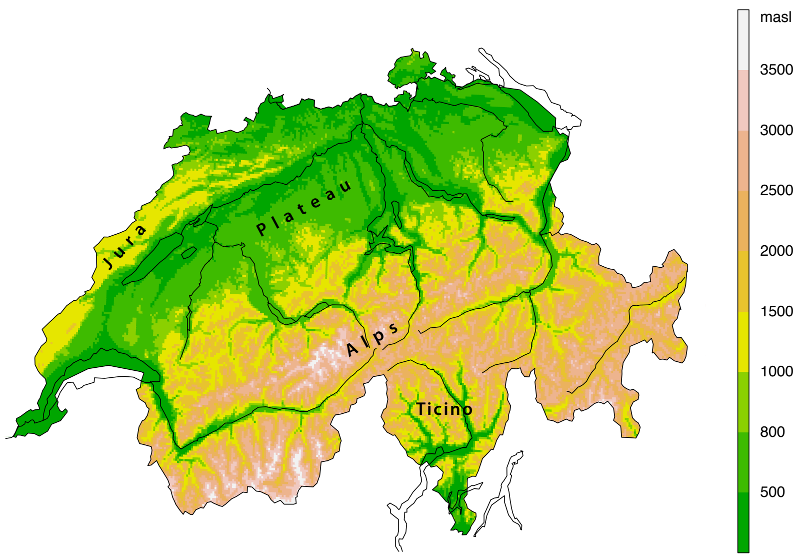

Figure 3.1: Digital elevation model of Switzerland, with the regions referred to in this study.

daily precipitation totals and mean daily temperature measurements. The number of stations used for each day of the analyses varies over time. For the time period considered here, it is around 430 to 520 stations for precipitation and 86 to 91 for temperature. Further details on the interpolation procedure can be found in Frei and Schär (1998), Schwarb (2000), Frei et al. (2006, section 4.1) for RhiresD and in Frei (2013) for TabsD. We upscaled (averaged) those 2 km datasets to the 25 km grid used for the RCM investigated in this study, and used the upscaled fields for the RCM evaluation. In continuation we refer to these upscaled fields as MS.

3.2.4 The challenging quantification of interpolation errors

Gridded observational data sets are inevitably subject to different kinds of errors, separated here in two categories. First, measurement errors: stations observations do not capture to the 'true' atmospheric conditions, as a result of the limited accuracy and precision of the measuring instruments and of deficiencies in the measurement process, such as the undercatch of precipitation in snowy or windy conditions. Second, interpolation errors: the interpolation process leads to different errors depending on the realism of the assumptions on which it relies and on the station density. In this study, as in the large majority of similar studies, measurement errors are not taken into account or corrected, and only interpolation errors are assessed.

When evaluating RCM outputs, a common way to account for observation uncertainty is to consider several observational data sets (e.g., Bellprat et al., 2012; Sunyer et al., 2013). It is important to stress that the spread among the considered data sets is not necessarily representative of the observational uncertainty. Most often, these data sets are not independent but share observations or rely on similar geostatistical principles. Hence a low dispersion of the ensemble is not a guarantee that the observational uncertainty is low. To hedge against this risk, a high number of observational data sets can be considered (e.g., Greve et al., 2014). Yet, a key issue that remains when several data sets are used, is that the reliability of these uncertainty estimates is typically unknown, i.e., the probability of the 'true value' of the observations to be within the spread of the considered data sets is unknown.

An alternative approach is to focus on a single data set and to quantify its errors. The efforts that went into the creation of the regional data set E-OBS (Haylock et al., 2008) and the global data set HadCRUT4 (Morice et al., 2012) are examples of such an approach. While HadCRUT4 resolution is too coarse for regional analyses such as the present one, E-OBS data and their associated error estimates were computed on the same grid as several ENSEMBLES climate projections (van der Linden and Mitchell, 2009) and were involved in a number of RCM evaluation studies (e.g., Christensen et al., 2008; Buser et al., 2009; Themeßl et al., 2011a; Kotlarski et al., 2014). The error estimates accompanying E-OBS-data, usually ignored in RCM evaluation studies, consist of the standard deviation of the interpolation error under the assumption of no systematic error. It is noteworthy that those estimates describe random errors, which often become insignificant for large sample sizes, for instance when averages over 30-year periods are computed (see e.g., Eqs. 1 and 2 in van der Schrier et al., 2013). In contrast, systematic errors are not averaged out over long periods (hence are those relevant for this study), but they were not quantified for E-OBS, as the precipitation and temperature estimates are assumed to be unbiased. Further, later analyses by the authors of E-OBS revealed that the error estimates are underestimated for both temperature and precipitation (Hofstra et al., 2009). They derived 95% confidence intervals based on their error estimates, but it appeared that a much larger fraction than the theoretical 5% of the observations fell outside of these intervals. We hence decided not to use E-OBS data and their error estimates.

Instead, we used estimates of the systematic error in monthly means of the RhiresM and TabsM datasets (the monthly equivalents of the daily RhiresD and TabsD), which were determined by leave-one-out cross-validations (MeteoSwiss, 2013a,b). Relative standard errors of +/− 20% in point estimates for Jura and the Plateau and of +/− 25-30% for the Alps and the Alpine south side were reported, and showed little seasonal variation (MeteoSwiss, 2013a). Note that given a mean precipitation amount over all stations for DJF and JJA of 2.9 and 4.4 and mm/day respectively, a 25% error corresponds to about 0.7 and 1.1 mm/day. For temperature, the reported standard errors in monthly means, also estimated by leave-one-out cross-validations, vary between 0.6°C (0.5°C) over the Plateau and Jura to 1.8°C (0.7°C) in the Alps and in Ticino in winter (summer, MeteoSwiss, 2013b). For both precipitation and temperature, the reported errors are point estimates, but in absence of estimates for larger areas, we use them here as (upper) estimates of the error in mean seasonal precipitation and temperature for the MS gridded data (Figure 3.6).

Errors are also introduced when interpolating gridded RCM simulations to station locations. Such interpolations are common in impact studies, when the impact model was calibrated with station data or when it requires a finer resolution than that of the RCM. We hence conducted another attempt to quantify interpolation uncertainty, by using a basic interpolation method. We estimated the systematic error introduced by using an inverse distance weighting (IDW) interpolation from a 25 km grid to station locations by cross-validation. For each station, the observed value was compared to the interpolated value obtained using the four closest stations out of subset of 70 stations chosen randomly from the 392 rain gauges or 86 temperature stations. The number of stations available for the interpolation was restricted to 70 to simulate the density of RCM grid points, as 70 of them fall within the boundaries of Switzerland. Although the randomly selected stations are not on a grid, their density is close to that of the grid points. Further details on the IDW method used throughout this study are provided in Section 3.2.5.

In continuation, we consider that when the RCM-observation difference is comparable to or larger than the estimated interpolation error (i.e., $|D/I| \geq 1$), then D is too large to be explained by interpolation errors alone. In contrast, for $|D/I| < 1$, the RCM-observation difference could simply be a consequence of the interpolation errors and it is unclear how large the model bias is.

3.2.5 Regional climate simulations and their interpolation

This study uses the climate simulations from the community land model (CLM) driven by the GCM HadCM3. The model run was performed by ETH in the framework of the ENSEMBLES project (van der Linden and Mitchell, 2009). When compared to E-OBS, this GCM-RCM chain was reported to overestimate winter precipitation and underestimate winter temperature over the Alpine range (Themeßl et al., 2011a). Interpolation from grid points to station locations was performed using inverse distance weighting, assuming a precipitation increase of 5% per 100m and a lapse rate of 0.5°C per 100m, as the commonly used value of 0.65°C per 100m appeared to be too high in our area and in other mountainous ones (Blandford et al., 2008; Immerzeel et al., 2014). IDW is a basic interpolation method with clear shortcomings. The resulting fields are unrealistically smooth, and usually poorly reflect local conditions, such as complex lapse rates or the anisotropy of precipitation fields influenced by the dominant wind (Tobin et al., 2011). These limitations lead to errors in the IDW-interpolated fields, which we quantify using cross-validation, and refer to in continuation as interpolation errors. Another limitation of IDW interpolation is its inability to disaggregate precipitation. Spatial disaggregation is important as it allows for the representation of severe and localized (sub-grid) precipitation events, whose intensity is higher than the mean intensity of the grid cell (Perica and Foufoula-Georgiou, 1996). This aspect is hence crucial when short periods or extreme events are assessed, but for this study focusing on 30-year means, we assume that the effects of the absence of disaggregation can be neglected.

The RCM-observation differences for 30-year seasonal means can reach a few °C or mm/day. To put these differences in perspective, we compared model simulations to the distribution of 30-year means estimated using bootstrapping. We constructed 100 time series of 30 years by sampling individual years with replacement from the precipitation and temperature records. We then considered the rank of the model simulation among these 100 time series. A rank of 1 (101) means that the average season is colder (warmer) in the model than in the coldest (warmest) 30-year time series produced by bootstrapping, thereby indicating a clear cold (warm) bias. Ranks closer to 50 indicate more realistic simulations, in which the simulated average is closer to the observed median.

3.2.6 Natural variability at the multi-decadal time scale

Internal variability is a crucial quantity, as it defines the lowest threshold to which the uncertainty in climate projections could be reduced, i.e., the remaining uncertainty in case of perfect models and a perfect knowledge of future greenhouse gas emissions. It has therefore been investigated by a number of studies with different purposes, ranging for the quantification of irreducible uncertainties to their importance for impact assessments (Hawkins and Sutton, 2009; Deser et al., 2012a; Mahlstein et al., 2012; Fatichi et al., 2014). There is no perfect way of quantifying natural variability. Instead, several competing approaches exist. Here we estimated natural variations in three different ways: using homogenized time series of precipitation and temperature, multiple runs of the same climate model and a boot-strapping approach. The three methods are discussed in continuation. We estimated natural variations over 30-year periods to reflect the length of our study period.

3.2.7 Estimating natural variability using homogenized time series

The first approach to estimate N follows Hawkins and Sutton (2009), referred to as HS09 in continuation. Firstly, the long-term forced signal in a measurement record is estimated by fitting a 4th order polynomial. This trend is assumed to represent the climate response to the external anthropogenic and natural forcing. Secondly, this forced component is removed from the time series. The fluctuations of the detrended time series are assumed to stem from the natural variability alone. Thirdly, a 30-year moving average is applied to the residual time series. N is finally estimated by computing the standard deviation of the residual time series. As Begert et al. (2005) stressed that trend estimation can be misleading if non-homogenized data are used, this analysis was limited to stations for which homogenized time series are available. Further, to capture multi-decadal oscillations, the analysis was only carried out for stations with a long homogenized record, starting in 1864 or before. This corresponds to 11 temperature stations and 9 rain-gauges. Fischer et al. (2012) relied on the same method to estimate N in ENSEMBLES climate projections and in the homogenized time series also used in this study, but focused their analysis on five stations.

3.2.8 Estimating natural variability using multiple runs of the same RCM

The second method used to estimate natural variability relies on multiple climate simulations. The model experiment was performed with the NCAR-DOE Community Earth System Model CESM 1.0.4 forced with observed historical radiative forcings up to 2005 and by the representative concentration pathway RCP8.5 thereafter (Moss et al., 2010). The experiment consists of 21 simulations started in 1950 from slightly different atmospheric initial conditions (see Fischer et al., 2013; Fischer and Knutti, 2014, for details). The simulations were performed with the exact same model, forced with the same radiative forcings and share the same initial conditions in all components except for the atmosphere. The spread among the simulations hence only represents the natural variability of the system, similar to the setting used by Deser et al. (2012a). To gain insights into the natural variability at the regional scale, the CESM runs were downscaled over Europe using the regional climate model COSMO-CLM at a resolution of 0.44° ($\sim$50 km). COSMO-CLM was hence run 21 times over the 1950-2100 period, each time using another realization of CESM as boundary conditions. A 30-year moving average was applied to the simulation of temperature and precipitation of each RCM run. The standard deviation across the 21 RCM runs was then extracted for each year and the value for the year 1994 was considered for 1980-2009. It is noteworthy that applying a 30-year moving average leads to a significant smoothing of the variations and consequently, the spread among the 21 runs is reduced (Figure 3.2). This illustrates that the mean state of the climate system over 30-year periods is more consistent across the different RCM runs than the exact succession of dry/wet (or cold/warm) years, which is more subject to natural variability.

Note that the RCM used to estimate D is also CLM (Section 3.2.5). It is however driven by a different GCM (HadCM3 and not CESM) and it is a different version of CLM. Given the complex topography of the study area, we decided to focus on the model with the finer grid and hence used HadCM3-CLM instead of CESM-CLM to compute D.

3.2.9 Estimating natural variability using boot-strapping

The third method used to estimate natural variability is based on boot-strapping and can be seen as the poor man's alternative to the methods requiring long homogenized time series or multiple climate model runs. It is inspired by the studies of Prudhomme and Davies (2008) and Bosshard et al. (2011). Individual years from the reference period 1980-2009 were resampled with replacement to create 500 time-series of 30 years. Each constructed time series usually contains several repetitions of the same observation year, leading to fluctuations of the 30-year mean and to spread among the 500 time series. For instance, if dry years were drawn more frequently than wet years, the resulting time series will present overall drier conditions than the original one. Using temperature and precipitation measurements from four stations spread over Switzerland, we found that the number of 500 constructed time series yields reasonably robust estimates of N without causing overly long computing times.

As a summary, the three methods used to estimate N rely on different periods (1864-2011, 1950-2100 and 1980-2009, respectively) and sample the variability in different ways (by considering 148 years of homogenized records, by running 21 climate simulations or by constructing 500 time series by bootstrapping, respectively).

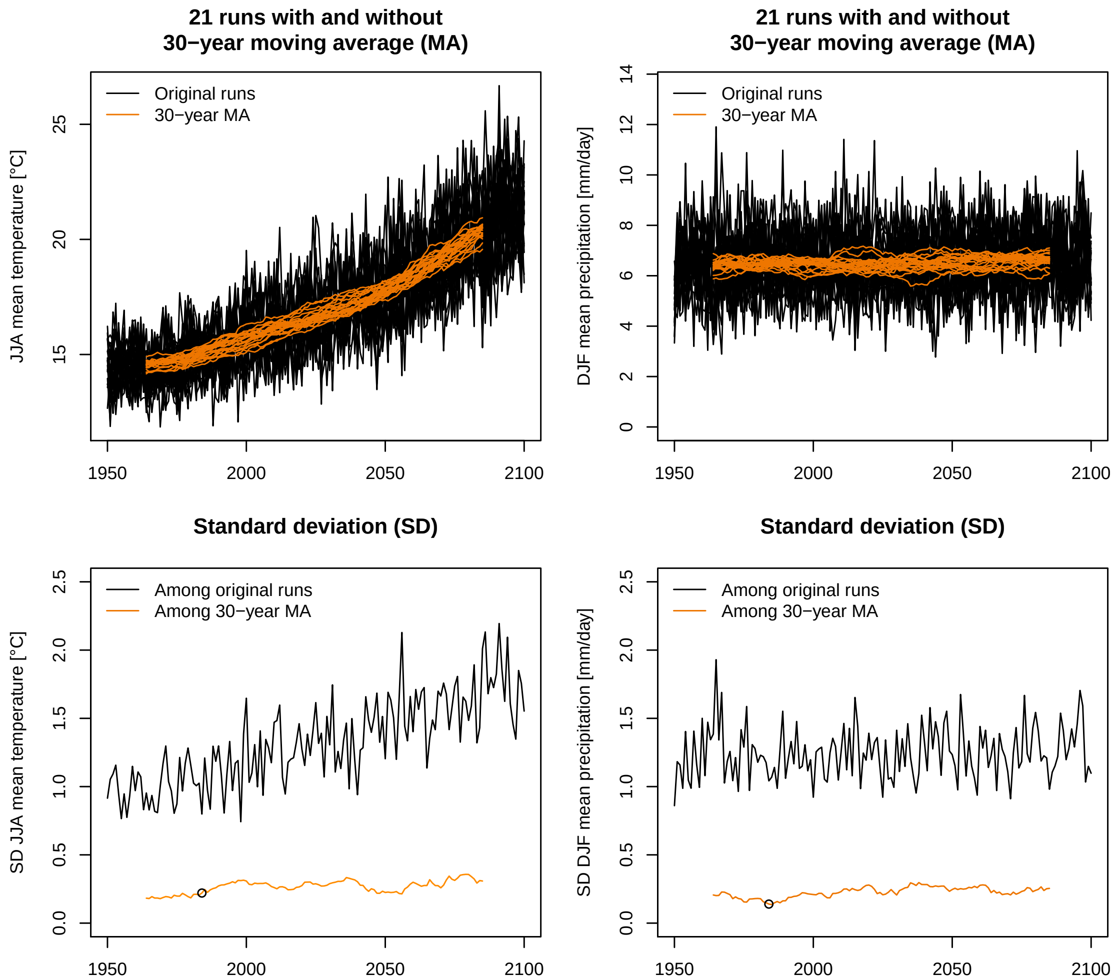

Figure 3.2: Illustration of the estimation of natural variability using 21 CESM-COSMO-CLM simulations and of its dependence on the time scale. The seasonal means of JJA temperature (left) and DJF precipitation (right) for one grid point on the Swiss plateau are shown (top) with their standard deviation (bottom). Focusing on the multi-decadal time scale (applying a 30-year moving average) reduces the spread among the simulations (top) and the estimated natural variability (bottom). The black circle (bottom) indicates the value reported for 1980-2009. Although CESM-COSMO-CLM is not evaluated here, note that the simulated DJF precipitation is significantly higher than the observations.

3.3 Results

3.3.1 Differences between observations and RCM simulations

The two most salient discrepancies between observations and RCM simulations, which appear both at the station and grid scale (Figure 3.3), are an overestimation of the mean precipitation amounts on the northern flank of the Alps in winter and an underestimation of winter temperature in the Alps in winter (Themeßl et al., 2011a). The wet winter bias is not specific to COSMO-CLM and has been reported in most of the ENSEMBLES GCM-RCM simulations for Switzerland (Fischer et al., 2012). This overestimation exceeded 2 mm/day over large areas. At those locations, the simulated DJF precipitation corresponds

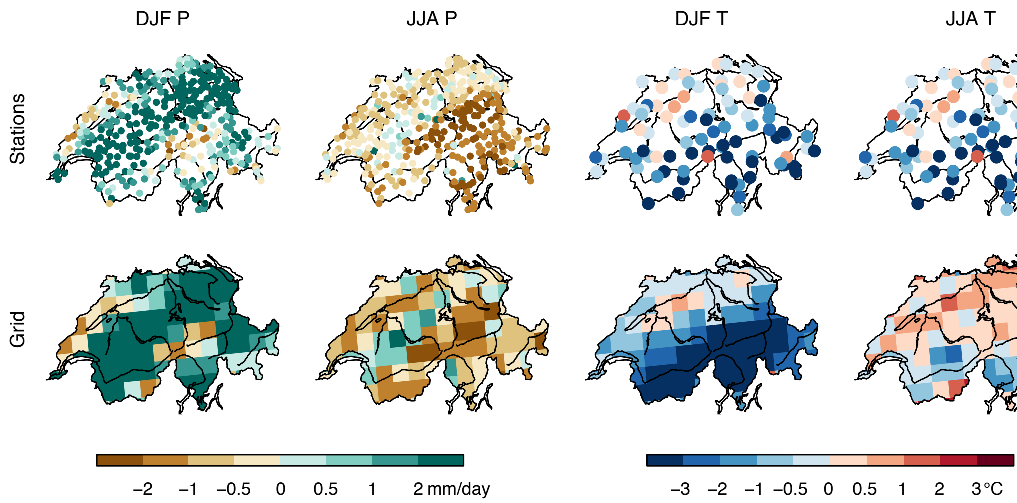

Figure 3.3: RCM-observations differences D. D was estimated comparing interpolated RCM simulations with station observations (first row, each dot corresponds to a station) and RCM simulations with the gridded MS data set (second row, on a 25 km grid). D is shown for winter and summer mean precipitation (first and second column) and for winter and summer mean temperature (third and fourth column). Note that the same column organization is used in Figures 3.4 to 3.7.

to more than twice the reference at grid and station scale. Overall, the absolute model-observation differences are larger in the Alps than at lower elevations and are larger in winter than in summer. Summer temperature seems to be underestimated at station scale, but not at grid scale. This probably stems from the IDW process, but regardless of the causes, this show that different data sets can lead to the different D and hence to situations in which D is ill-defined (e.g., Sunyer et al., 2013).

Figure 3.4 shows that over most of Switzerland, the average winter precipitation (temperature) simulated by the RCM is higher (lower) than that of every single 30-year time series produced by bootstrapping observations. In other words, the average simulated conditions correspond to extremely rare situations given the observations. Such deviations from the climatology will propagate if the climate simulation is used directly to force an impact model, resulting into artificially rare impacts. The transects across the Alps show that in winter, precipitation and temperature simulated by the RCM clearly fall outside of the envelope of time series obtained by bootstrapping. The overestimation of winter precipitation on the northern flanks of the Alps and the Swiss plateau (latitude $> 47°$ North) and the underestimation of winter temperature in the Alpine range (latitude $\sim 46.6°$ North) are particularly severe. Summer conditions are better captured, but in large parts of the country, the departure of the simulations from the observations is larger than the natural fluctuations estimated by bootstrapping.

3.3.2 Differences in natural variability estimates

Natural variability estimates for 30-year means in Switzerland are typically in the range 0.05-0.30 mm/day for precipitation and 0.10-0.35°C for temperature (Figure 3.5). They are consistent with those reported by Fischer et al. (2012). Overall there is a good agreement on the order of magnitude of N among the different methods. However, there are differences

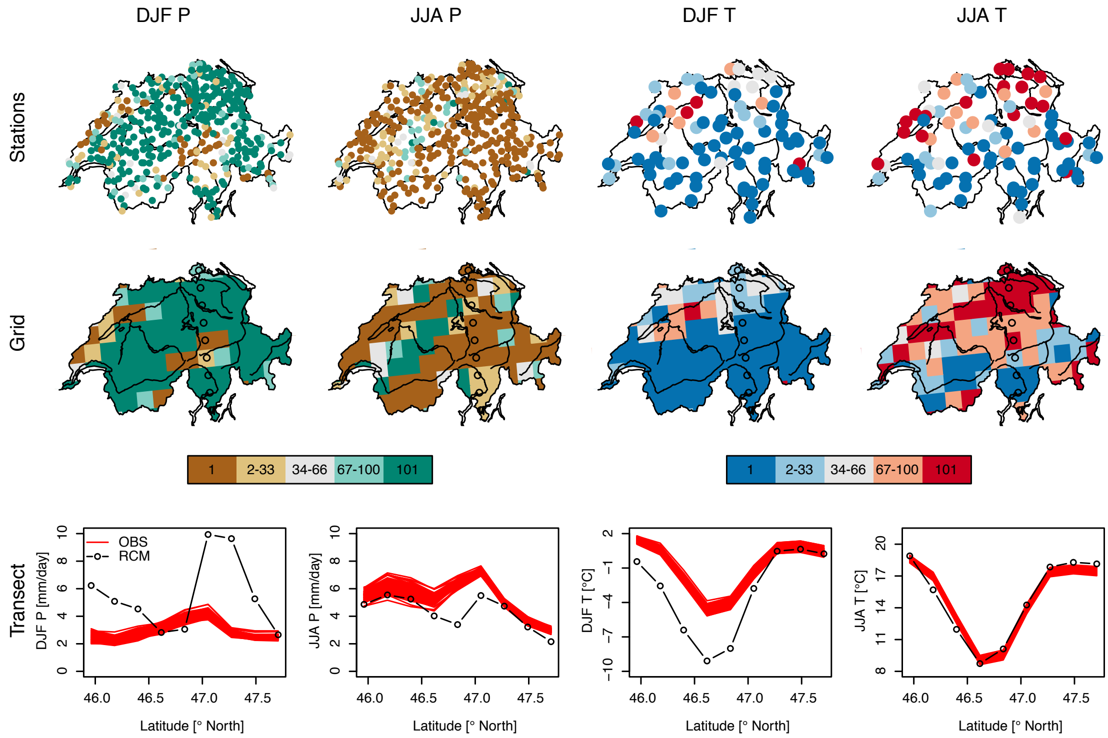

Figure 3.4: Rank of the simulated 1980-2009 seasonal mean among the seasonal mean of the 100 time series constructed by bootstrapping (top and middle row). A rank of 1 or 101 indicates that the simulated average season is outside of the envelope obtained by bootstrapping. The open circles in the middle row indicate grid cells used to for the transects of precipitation and temperature across the Alps shown in the bottom row.

in the temporal and spatial variations of N.

The different methods yield different seasonal variations of N, especially for temperature. N in temperature is larger in winter than in summer when homogenized observations for 1864-2011 are used, owing to the stronger variability on the synoptic scale during this season (Fischer et al., 2012). The same holds when bootstrapping is used. Yet, when the standard deviation among the 21 RCM runs is used, the natural variability of temperature is higher during summer. We suspect that N is somewhat overestimated as it has been demonstrated for interannual variability in summer for the COMSO-CLM and other regional climate models (Vidale et al., 2007; Fischer and Schär, 2009). This stresses that the results based on RCM experiments are conditioned on the model ability to capture natural variability. For precipitation, each method delivers comparable estimates for winter and summer. Yet discrepancies appear among the different methods, with lower and higher estimates delivered by HS09 and multiple RCM runs, respectively.

Overall, the spread among the RCM runs yields higher estimates than the bootstrapping method and HS09 approach. Bootstrapping based on 30 years of record cannot fully account for multi-decadal oscillations and consequently is likely to underestimate N. Similarly, there is a risk that the fourth-order polynomial used in HS09 filters some of the multi-decadal variability and thus that the estimates of N are biased low. To assess this, we used a linear regression instead of a fourth-order polynomial to estimate the long-term forced signal. For summer temperature, this leads to an increase of the estimated natural

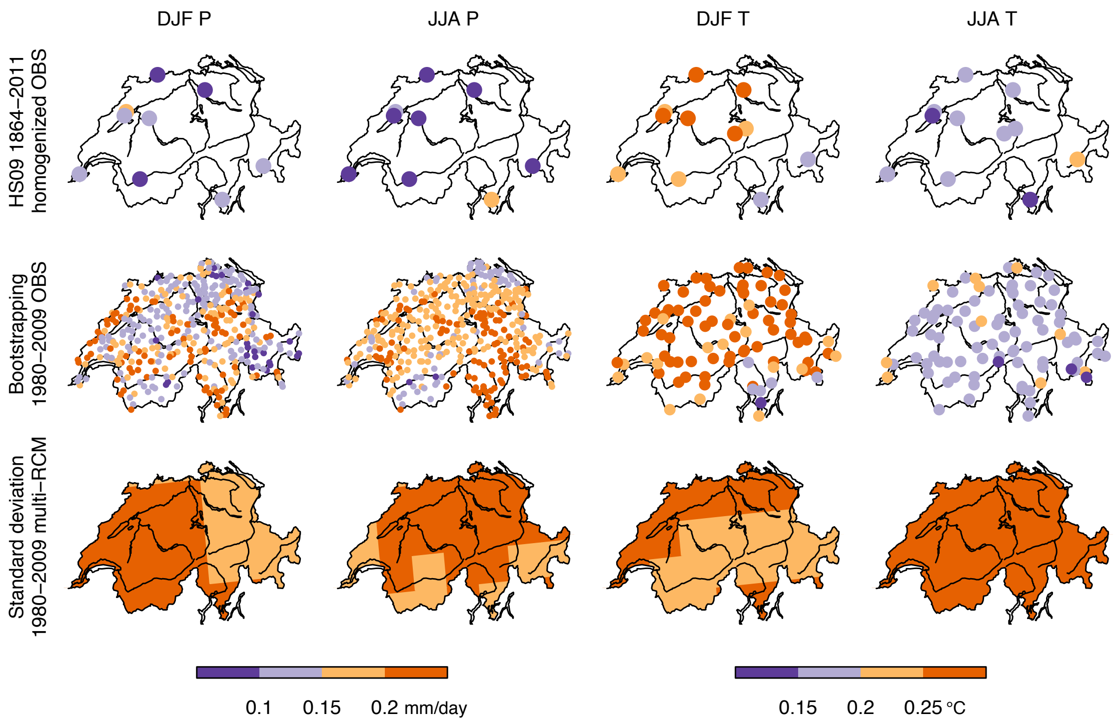

Figure 3.5: Natural variability N estimated using HS09, bootstrapping and multiple RCM runs (top to bottom row). For the two top rows, the number of stations depends on the availability of observations for the period considered.

variability of 54% (average over all the stations). For winter temperature and precipitation, the increase is weaker and is of 10% or less. So natural variability is overall lower when computed using HS09 than with the other two methods, except when it is estimated for summer temperature using a linear regression instead of a fourth-order polynomial.

The different methods also lead to different spatial patterns of N. For instance, when station observations are used, the Ticino region is associated to a larger N for precipitation, but this pattern does not appear in the estimates derived from RCM simulations, as its 50 km grid is arguably too coarse to capture these spatial variations. This is supported by the range of values estimated using RCM data, which is narrower than when station data are used. Yet, when analyzing the estimates based on stations data, there is no consistent spatial pattern or a clear relation between N and stations elevation. Rather, the estimated natural variability appears to be more sensitive to the method used and the input data, than to the station location.

Overall, our analysis reveals uncertainties in the estimates of natural variability on multi-decadal time scales. Although some clear features emerge, the different methods and data sets produce divergent estimates, to an extent that sometimes makes unclear which season or which region shows the highest or lowest natural variability in Switzerland. Further, N also varies with time (Figure 3.2), which represents yet another source of uncertainty in its quantification.

3.3.3 Interpolation errors

To investigate seasonal and spatial patterns in interpolation errors, cross-validation was applied to two data sets: IDW interpolated data and TabsM and RhiresM gridded data. Because of the difficulty to account for the dependence of these error estimates on the spatial and temporal scales (see our assumptions in Section 3.2.4), these estimates should be interpreted with caution. Yet some clear patterns emerge. Irrespectively of the season, parameter or data set, interpolation errors are generally higher in Alps than on the Plateau or the Jura. A station can be subject to a strong overestimation, while a neighbouring station is undermined by a strong underestimation, indicating that the interpolation process does not fully capture local climatic conditions. This is particularly clear for winter temperature in the Alps, and although more advanced methods than IDW reduce this issue, they do not completely solve it (Figure 11 in Frei, 2013). These changes in sign over distances often smaller than the RCM grid spacing lead to interpolation errors that tend to be smaller at the grid scale than at the station scale.

For temperature, errors are larger in winter, reflecting the limited ability of the interpolation schemes to capture high spatial heterogeneity, related in particular to winter cold pools in alpine valleys. For IDW, the absolute differences between the interpolated and measured 30-year seasonal mean are 1.2°C in winter and 0.7°C in summer, when averaged over all the stations. For precipitation, the IDW absolute error in 30-year seasonal means is about 0.5mm/day in both DJF and JJA when averaged over all the stations. These typical errors are larger when computed over shorter periods of time (e.g., 10 years), which illustrates the dependence of observational uncertainty estimates on the temporal scale.

Comparison with observational uncertainties estimates from other studies is difficult, because of the unique combination of climate, topography, station density, temporal and spatial scales embedded in each study. As an example, Isotta et al. (2014) characterized both systematic and random errors for a 5 km pan-Alpine gridded precipitation data set for different precipitation intensities. For an alpine area with dense network, they for instance report over- and underestimations of daily precipitation events of moderate intensity by 25% or more in half of the cases. They however do not estimate typical errors at the monthly or seasonal timescales.

3.4 Discussion

3.4.1 The influence of natural variability on the multi-decadal time scale

We argue that our ability to evaluate and bias-correct climate projections is compromised if natural variability, which is an unpredictable quantity after a decade at most, explains a significant part of the RCM-observation differences. We see four main reasons therefore. Firstly, it would mean that significant fluctuations in D can be expected from one 30-year period to the other, which goes against the assumption of bias stationarity over time, on which most bias correction studies rely (see also Chen et al., 2011; Maraun, 2012; Chen et al., 2015). Secondly, as these changes originate from the chaotic nature of the atmosphere and cannot be predicted over several decades, they may not be corrected in a deterministic way. Thirdly, rerunning the same model from different initial conditions could lead to a significantly different D over the reference period, hence to a different skill, which should imply that robust model evaluation should be performed using an ensemble of simulations of the same model sampling the natural variability. Finally, a large contribution of the natural variability to D would challenge the assumption that a model performing well

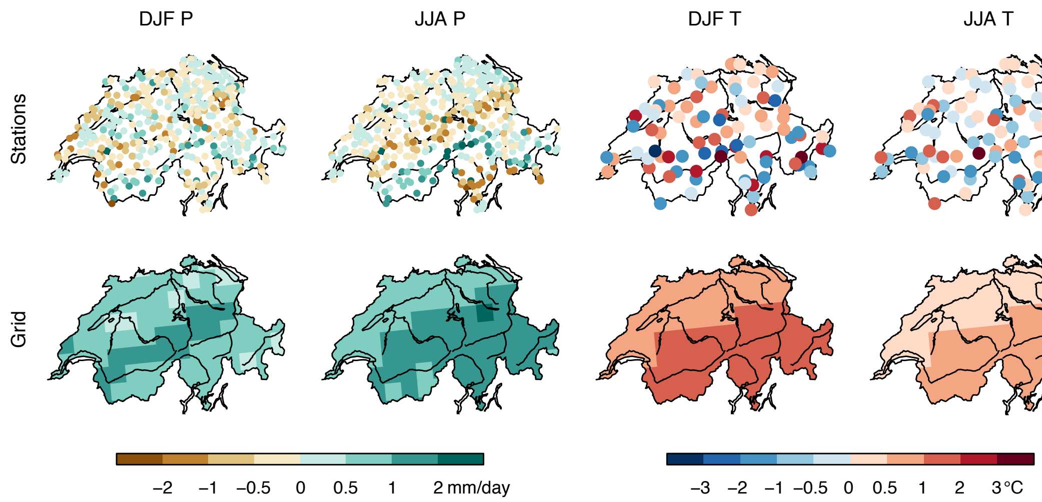

Figure 3.6: Interpolation errors I estimated using cross-validation of IDW interpolation (top row) and the estimated standard error (> 0) in TabsM and RhiresM (bottom row).

under present conditions will also perform well in the future, since the result of its good performance during the reference period might simply be the result of a favorable natural variability state. Note nevertheless, that even in absence of natural variability, the constant bias assumption can be violated by state-dependent model errors.

To explore the contribution of natural variability to RCM-observation differences we normalized these differences by the estimated natural variability. At grid scale, RCM simulations are compared to MS and the natural variability is derived from the standard deviation among the multiple RCM runs on a 50 km grid and then interpolated on the 25 km grid using a bilinear interpolation (the interpolation step should not alter our conclusions; Rauscher et al., 2010). At station locations, RCM simulations interpolated using IDW are compared to station observations and the natural variability is derived by bootstrapping. In both cases, we consider that fluctuations stemming from natural variability are normally distributed with a standard deviation N. We hence reject the hypothesis that D stems from natural variability alone when $|D/N| > 2$, which corresponds to an offset greater than 2 standard deviations and occurs with a frequency of about 5% in normal conditions. Since D can be negative, we consider the absolute value of D/N to focus on the emergence of the model biases independently of their sign.

In few cases, D is within the range of the natural variability (gray areas in the two top rows of Figure 3.7). It is especially the case for JJA temperature at grid scale, but we attribute this to the suspected overestimation of the N for JJA temperature, which artificially lowers the D/N ratio. For most locations, D is too large to be explained by the natural variability alone (green areas). In those cases, rerunning the climate model with different initial conditions is unlikely to lead to significant changes in D. That is, the risk of the difference changing over time as a result of the natural variability is low. This is supported by the median of the D/N ratios reaching 3.6 to 13.3 when computed over the country (Figure 3.7, we do not consider the value of 1.4 obtained at grid scale for JJA temperature, as it is probably biased low), which emphasizes that N is significantly smaller than D.

We used different approaches to estimate the natural variability and showed that un-

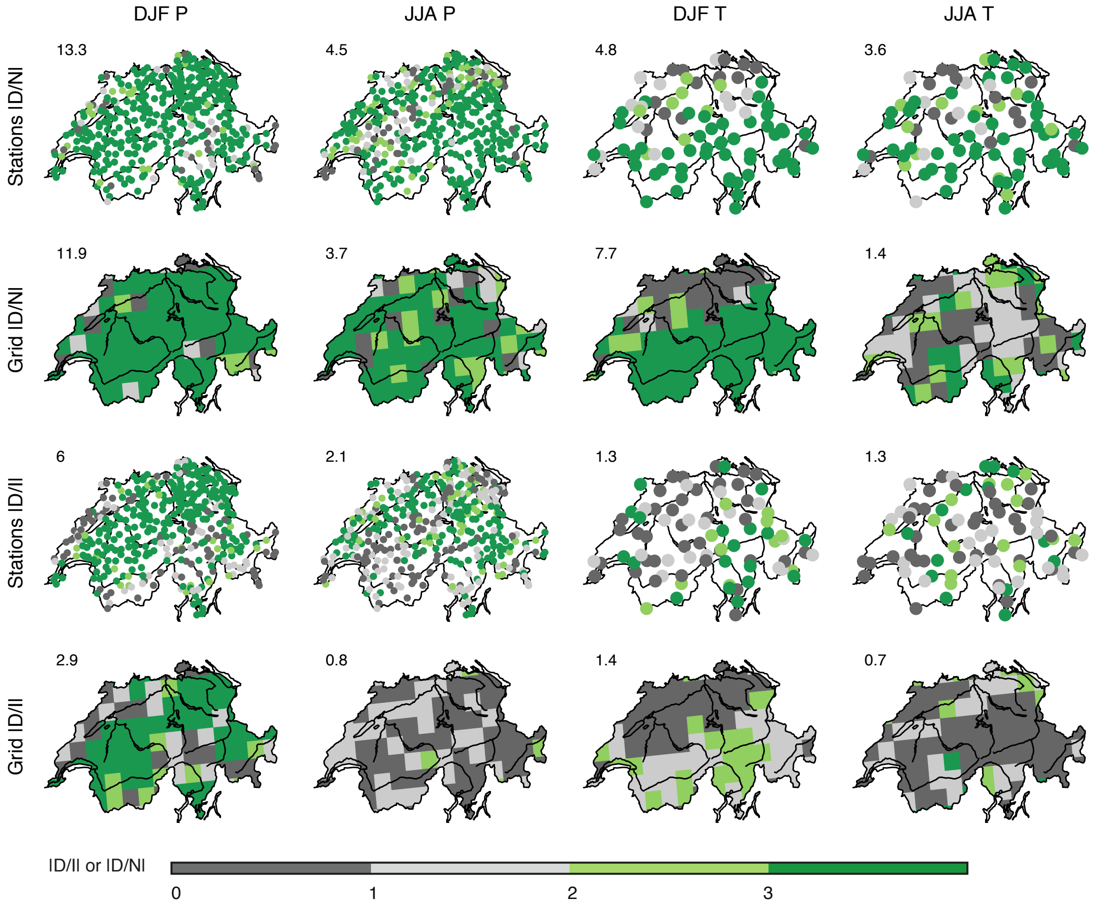

Figure 3.7: Comparison of the RCM-observation differences to the estimated natural variability (two top rows) and the estimated interpolation errors (two bottom rows). The median value of the $|D/N|$ or $|D/I|$ ratios over Switzerland is indicated in the top left corner of each map.

certainties in its seasonal and geographic patterns are induced by the methods (e.g., bootstrapping or HS09), by the data (e.g., homogenized observations or 30-year of observations) and by the eventual reliance on a climate model. Yet, we argue that, in our study area, the discrepancies between RCM simulations and observations over 30-year period are most often too large to be stem from natural variability alone, despite uncertainties in the exact value of N.

However, we expect this result to be conditional on the multi-decadal time scale and the mean quantities that this study focuses on. If the analysis were performed on a shorter time scale, the natural variability would be larger (Figure 3.2) and explain a larger portion of the RCM-observation mismatch. Further, although mean quantities are central characteristics from impact assessments, as they set the baseline of the modelled system, they are less subject to natural variability than other variables. In particular when the focus is on trends, or on changes in the intensity and frequency of extreme events, natural climate variability plays a more significant role and needs to be accounted for (Perkins and Fischer, 2013). Piani et al. (2010) for instance showed that the bias-correction applied to GCM simulations of precipitation and temperature varies with time when consecutive 10-year periods are considered, and also varies with the initial conditions of the GCM, as a result of natural variability.

3.4.2 The influence of interpolation errors

As in the previous section, two sets of data were analysed, one at station scale, based on the interpolated RCM using IDW, and one at grid scale relying on MS as observational data set. Typical interpolation errors were estimated by cross-validation. Figure 3.7 shows that, in contrast to N, I is often as large as D. This is shown by the larger extent of the gray areas in the two bottom than in the two top rows. For instance, for DJF temperature, D is too large to be explained by N alone (large green areas), yet the same D is comparable to I (large gray areas) and the model bias emerges from interpolation errors only in some parts of the Alps. We infer that comparing the same model to different observational data sets produces larger changes in D than rerunning the same model from different initial conditions. In other words, we find the RCM-observation mismatch over 30-year periods identified here depends significantly on interpolation errors, whereas multi-decadal natural variability only accounts for a minor fraction of it.

We propose that in the gray areas of the $|D/I|$ maps, it is unclear whether the model is biased or not. First, in those cases with small RCM-observation differences, the model can still be biased if this bias is compensated by I and N. Second, even if there is an RCM-observation mismatch, it might still be within the uncertainty range of the observations. For instance for summer temperature and precipitation at grid scale, as the interpolation error is comparable to the RCM-observation differences, we argue that referring to this mismatch as 'model bias' is misleading. Conversely, when D is significantly larger than I and N, independently of the method and data used to estimate these factors, then we argue that there is a clear emergence from the model bias. The clearest emergences are the wet bias on the northern flank of the Alps and, to a lesser extent, the cold bias in the central Alpine valleys, both in winter.

However, beside these two clearly emerging biases, there is no clear boundary between areas in which the model bias emerges and those where it does not. This is in particular related to the difficulty to quantify interpolation errors. Clearly, in this study, the lower the estimated interpolation error I, the more likely a given bias in the model is to emerge from it. For instance, for summer precipitation, the interpolation errors estimated for IDW tend to be smaller than for MS (Figure 3.6). It follows that the model biases emerge using station data but not when MS is used. Because both methods used here to estimate I present clear limitations, it is unclear where the biases emerge from interpolation errors. Yet, the approach outlined here illustrates that a better observational data set, i.e., with lower errors quantified in a more reliable way, would enable a more detailed and nuanced evaluation of model performance.

Major efforts have gone into downscaling techniques disaggregating precipitation in time and space. Here we use the simple IDW technique as a benchmark, producing larger error than more advanced methods such as kriging (Tobin et al., 2011). Yet, the errors introduced by this basic method are in some cases still smaller than the biases in RCM simulations, which underscores the importance of downscaling methods performing both disaggregation and bias correction.

3.4.3 Toward a more systematic consideration of observational uncertainties

Because climate models play such a pivotal role in our attempts to understand and anticipate future changes, there is a strong need for data sets to diagnose their performance in a robust way. While climate simulations and projections are now generally considered in an ensemble way, observations are still overwhelmingly handled in a deterministic way. This study and others (e.g., Gómez-Navarro et al., 2012; Sunyer et al., 2013) highlight the need for a more systematic consideration of uncertainties in the observational data sets when evaluating and bias-correcting climate simulations.

Comparing the same RCM simulation to different observational data sets can lead to different estimates of D, which implies that biases are not as clearly defined as perceived when only one observational data set is taken into account. Thus, evaluations relying on a single observational data set with no assessment of its uncertainty need to be interpreted with caution. Similarly, as observational data sets are not error-free, care must be used when involving them in the bias correction the RCM simulations. When a data set is heavily biased, a method like quantile mapping (Piani et al., 2010; Teutschbein and Seibert, 2012) could even result in moving the cumulative distribution function into the wrong direction, thereby enhancing the error.

One solution to these issues is to evaluate the robustness of skill and biases across a wide range of observational data sets. The robustness of the model performance and of its bias across the data sets could be depicted using stippling and hatching on top of a map representing the mean performance or bias (e.g., Knutti and Sedláček, 2012; Tebaldi et al., 2011). In the context of climate impact modeling using bias-corrected climate simulation, a way to deal with the uncertainty in bias characterization would be to propagate it through the impact model chain. That is, running the same impact modeling chain using different observational data sets for bias characterization. Similar approaches have been already applied to explore the consequence of bias correction on hydrological modeling (e.g., Cloke et al., 2012; Muerth et al., 2013), but to our knowledge, not to account for the uncertainty in the observational data sets used for bias correction. We propose to include the observational data set as an additional source of uncertainty, and rely on an ensemble of bias correction techniques, based on different assumptions on the future bias evolution and on several observational data sets. This would result in an additional element in the cascade of uncertainty depicted in Wilby and Dessai (2010) and in its further widening. Further, if shorter time scales were considered, for which natural variability is expected to play a more significant role, a similar approach could be adopted to evaluate the sensitivity of the bias-corrected time series to these variations.

Another solution would be based on the use of a data set with reliable error estimates. When using the spread among data sets to quantify the observational error, each data set has usually the same weight and there is not attempt to distinguish between better and worse data sets. Hence if a better (with lower errors) quality data set became available, it would be merged with the lower-quality data sets, probably without changing significantly the perceived observational uncertainty. In contrast, using a single data set with reliable probabilistic error estimates would allow characterizing biases and evaluating models in a sounder way.

3.5 Conclusions

Interpolation errors, natural variability and model biases all contribute to the RCM-observation mismatch. Here we relied on several data sets and methods to estimate their respective contributions and discussed the consequences for model evaluation and bias correction.

Significant differences exist between RCM simulations and observations, both at grid and station scale. In winter, in large parts of the Alpine domain, the mean simulated conditions are so wet or so cold, that they correspond to extremely rare situations given the observations. These large deviations from the climatology impede the direct use of RCM simulations for impact modeling. To mitigate their effects over long simulation periods, it is key to investigate whether they are robust despite natural variability and uncertainties in observational data sets.

Natural variability is a crucial factor as it can potentially lead to important non-stationaries of the RCM-observations differences, but it is not predictable beyond a few decades. Different quantification methods were used, and although they delivered different estimates of the seasonal and spatial variations, they all indicate that, on the multi-decadal time scale, natural variations are considerably smaller than RCM-observation differences. It follows that rerunning the climate model with other initial conditions will probably not lead to significant changes in the RCM-observation mismatch. Yet, we expect that if the period of interest was shorter than 30 years or if instead of the 30-year means, e.g., high precipitation percentiles were considered, the role of internal variability would be larger and may easily account for a significant proportion of the RCM-observation differences.

Further, we find that RCM-observations differences often fall within the uncertainty range of the observation data sets. It means that in those areas, mismatches between RCM simulations and observations do not come from systematic errors in the RCM alone, but instead may stem to a large extent from interpolation errors. It follows that using different observational data sets for the evaluation and bias correction of climate simulations will lead to different assessment of their skill and to different post-processed time series. This stresses the importance to account for observational uncertainties, by for instance propagating them through bias-correction methods and impact models.

Finally, some model biases clearly emerge from natural variability and interpolation errors, independently of the method and data chosen to quantify these variables. However, the boundaries of the regions in which biases emerge are still fuzzy, because of the difficulty to produce estimates of observational uncertainties covering the wide range of temporal and spatial scales involved in climate modeling. This highlights the need for more systematic uncertainty assessments released together with future observational data sets. Overall we conclude that a better understanding of the reasons behind the RCM-observations mismatch under present conditions is essential to strengthen the interpretation of climate projections and to improve their use for impact assessments.

Acknowledgements

We thank the EU-FP6 project ENSEMBLES project for producing and providing the RCM simulations (available from http://ensemblesrt3.dmi.dk/) . We are grateful to the Swiss Federal Office of Meteorology and Climatology, MeteoSwiss for the station observations and the gridded data sets (data requests should be sent to MeteoSwiss). We acknowledge Urs Beyerle, Daniel Lüthi and Sven Kotlarski for their support in setting up the multimember RCM experiment (data requests should be sent to erich.fischer@env.ethz.ch) and Christoph Frei for providing the map plotting package. The constructive comments of two anonymous reviewers were greatly appreciated. The data analysis and visualization were performed with R (R Core Team, 2014). We thank the University of Oregon, USA, for hosting NA while working on this paper and the Swiss National Science Foundation for supporting this study (grant 200021_131995).

Chapter 4

Propagation of biases in climate models from the synoptic to the regional scale: implications for bias-adjustment

Nans Addor[1], Marco Rohrer[2], Reinhard Furrer[3], Jan Seibert[1,4]

[1] Department of Geography, University of Zurich, Zurich, Switzerland
[2] Oeschger Center for Climate Change Research and Institute of Geography, University of Berne, Berne, Switzerland
[3] Institute of Mathematics, University of Zurich, Zurich, Switzerland
[4] Department of Earth Sciences, Uppsala University, Uppsala, Sweden

Accepted pending revisions for publication in
Journal of Geophysical Research: Atmospheres

Abstract

Bias-adjustment methods usually do not account for the origins of biases in climate models and instead perform empirical adjustments. Biases in the synoptic circulation are for instance often overlooked when post-processing regional climate model (RCM) simulations driven by general circulation models (GCMs). Yet considering atmospheric circulation helps to establish links between the synoptic and the regional scale, and thereby provides insights into the physical processes leading to RCM biases. Here we investigate how synoptic circulation biases impact regional climate simulations, and influence our ability to mitigate biases in precipitation and temperature using quantile mapping. We considered 20 GCM-RCM combinations from the ENSEMBLES project and characterized the dominant atmospheric flow over the Alpine domain using circulation types. We report in particular a systematic overestimation of the frequency of westerly flow in winter. We show that it contributes to the generalized overestimation of winter precipitation over Switzerland and this wet regional bias can be reduced by improving the simulation of synoptic circulation. We also demonstrate that statistical bias-adjustment relying on quantile mapping is sensitive to circulation biases, which leads to compensation errors in the post-processed time series. Overall, decomposing GCM-RCM time series using circulation types

reveal connections missed by analyses relying on monthly or seasonal values. Our results underscore the necessity to better diagnose process misrepresentation in climate models, in order to progress with bias-adjustment and impact modeling.

4.1 Introduction

Climate model simulations can show large departures from observations. These biases are often subtracted and then forgotten, or post-processed for impact modeling using pragmatic methods that typically do not account for the origins of the biases. Yet to improve climate projections, design robust bias-adjustment methods and derive uncertainty estimates, a better understanding of the physical processes leading to biases is necessary. Biases are the result of accumulated errors in the representation of various processes occurring at various spatial and temporal scales. A direct and critical consequence is that biases are difficult to reduce. Further, biases are not stationary over time (Chen et al., 2015) and equally plausible assumptions on their evolution can lead to significantly different projected changes (Buser et al., 2009). In this article, we explore how a better diagnostic of the reasons behind biases can help with 1) the assessment of how much benefit could be gained by specific improvements (e.g. higher resolution in RCMs or better synoptic circulation in GCMs), 2) the better understanding of why biases remain after bias-adjustment and 3) the design of more robust bias-adjustment methods that are not purely empirical but account for known process misrepresentations. We therefore use the example of the propagation of errors from the synoptic to the regional scale in winter over an alpine region.

Biases in regional climate model (RCM) simulations emerge from processes at different spatial scales. At the regional scale, they were shown to stem from processes related to, for instance, soil moisture (e.g., Bellprat et al., 2013), surface albedo (e.g., Maraun, 2012) and convective precipitation (e.g., Ban et al., 2014). This study focuses on biases inherited from the larger synoptic scale, i.e., those biases stemming from the synoptic circulation represented by the driving model, typically a general circulation model (GCM, e.g., Noguer et al., 1998; van Ulden and van Oldenborgh, 2006). Considering the synoptic situation links the air flowing over the region of the interest to its origin, and thereby can deliver major insights into the quality of process representation by climate models (James, 2006; Maraun et al., 2011) and into the reasons behind the projected climate changes (e.g., Cattiaux et al., 2013).

To describe the synoptic circulation, a common approach is to rely on circulation types (CTs), as they enable to reduce a multivariate atmospheric flow to a few classes (Philipp et al., 2010). CTs are used routinely by weather forecasting services (e.g., Schiemann and Frei, 2010) and the identification of the synoptic situations causing extreme events is an area of active research (Stahl and Demuth, 1999; Prudhomme and Genevier, 2011; Wilby and Quinn, 2013). Once such a relation between CTs and extreme events is established, it is tempting to use it to infer future changes in extremes based on the CT changes projected by climate models. Yet an important first step is to assess how well climate models represent CTs under present climate (Huth et al., 2008). Several studies indeed reported biases in the frequency of CTs simulated by climate models, and found that they induce biases in regional surface variables, such as precipitation and temperature (van Ulden and van Oldenborgh, 2006; van Ulden et al., 2007; Blenkinsop et al., 2009; Plavcová and Kyselý, 2011).

Although the last years have seen the development of a profusion of bias-adjustment methods, in general these methods do not relate the biases to the (mis)representation of physical processes in models, such as the representation of the synoptic circulation, but instead perform an empirical adjustment. An example is the widely used quantile mapping method, which relies on the transformation of the RCM output so that, after the transformation, its cumulative distribution function matches that of the observations (Piani et al., 2010; Themeßl et al., 2011a). This method gained popularity in particular because of its ability to correct a wide range of statistical variables, such as the frequency of dry days or the variance of daily temperature (Teutschbein and Seibert, 2012). Closer scrutiny however reveals that other aspects crucial for impact modeling are not corrected correctly: the diurnal temperature range (Thrasher et al., 2012), the precipitation under different circulation types (Bárdossy and Pegram, 2011) and multi-day statistics (Addor and Seibert, 2014). Further, although quantile mapping has been shown to retain the relationship between temperature and precipitation simulated by climate models, recent studies suggest that it does not correct this relationship when it is biased (Wilcke et al., 2013; Li et al., 2014). Such issues may be fixed by further increasing the complexity of the statistical post-processing. . However, their recurrence reminds us that not identifying and accounting for the physical processes leading to the biases implies the risk of remaining biases in statistically post-processed simulations. It is hence fair to acknowledge that this kind of post-processing is not a 'correction' of the biases but rather an 'adjustment', and to refer to it as such.

To progress with the attribution of biases in RCM simulations, and the understanding of their consequence for bias-adjustment methods, this study explores how biases in synoptic circulation frequency propagate to the regional scale and influence our ability to bias-correct climate simulations. We address the following questions:

1. How well do GCM-RCM of the ENSEMBLES project (van der Linden and Mitchell, 2009) capture the frequency and regime of CTs in the Alpine region under present climate?

2. How does this influence biases in GCM-RCM simulations of mean temperature and precipitation over Switzerland?

3. What are the implications for bias-adjustment, in particular when quantile mapping calibrated for different CTs is used?

4.2 Data and methods

4.2.1 Climate models and evaluation procedure

This study was based on 20 GCM-RCM simulations produced for the ENSEMBLES project (van der Linden and Mitchell, 2009, see Supporting Table 4.1 for the model list). These GCM-RCM combinations were selected by Fischer et al. (2012) to produce the probabilistic Swiss Climate Change Scenarios CH2011 (CH2011, 2011). We evaluated each combination by extracting the grid points falling within the borders of Switzerland, an area of ∼41000 km^2. Model performance varies within this area, in particular with topography (e.g., Addor and Seibert, 2014), but here our focus is on the average performance over the whole country. We compared the simulated precipitation and temperature in winter (DJF) and summer (JJA) to gridded observations from the 2 km data sets RhiresD (precipitation) and TabsD (temperature Frei, 2013). These two data sets were recently released by the Swiss Federal Office of Meteorology and Climatology and take advantage of the dense network of stations across the country (∼420 daily measurement for precipitation and ∼90 for temperature). As a comparison, E-OBS 8.0 data set (Haylock et al., 2008), which is popular for RCM evaluation, relies on precipitation and temperature data from only 37 Swiss stations. Observed and simulated fields were averaged over Switzerland and then compared

An important aspect to take into account when comparing RCM simulations to gridded observations is observational uncertainties. Errors in observational datasets, which come in particular from the interpolation of station measurements to grid points and from measurement errors (e.g., precipitation undercatch), can be as large as the differences between RCM simulations and the gridded observations (Addor and Fischer, 2015). Their quantification can be estimated by comparing several datasets or by cross-validation, both methods presenting weaknesses. A source of error particularly relevant to this study is precipitation undercatch, which is particularly high in winter, when it is estimated to lead to an underestimation of precipitation by ∼8% below 600 masl and by ∼40% above 1500 masl (Frei et al., 2003). This is however not accounted for in the RhiresD data set used in this study, and to our knowledge there is at present no data set covering Switzerland on a regional grid that accounts for precipitation undercatch. In this study we assume that up to 30% of the mean precipitation over Switzerland can be missed by gauges in winter.

Switzerland was chosen as a study area because it constitutes an interesting test bed in

ID	Institute	GCM	RCM	Color in Figures
1	SMHI	BCM	RCA3	orange
2	METNO	BCM	HIRHAM	orange
3	ETHZ	HadCM3Q0	CLM	blue
4	HC	HadCM3Q0	HadRM3Q0	blue
5	UCLM	HadCM3Q0	PROMES	blue
6	VMGO	HadCM3Q0	RRCM	blue
7	METNO	HadCM3Q0	HIRHAM	blue
8	HC	HadCM3Q3	HadRM3Q3	blue
9	SMHI	HadCM3Q3	RCA3	blue
10	C41	HadCM3Q16	RCA3	blue
11	HC	HadCM3Q16	HadRM3Q16	blue
12	MPI	ECHAM5-r3	REMO	red
13*	DMI*	ECHAM5-r3*	HIRHAM*	*
14	KNMI	ECHAM5-r3	RACMO	red
15	SMHI	ECHAM5-r3	RCA3	red
16	ICTP	ECHAM5-r3	REGCM	red
17	CNRM	ARPEGE	ALADIN	green
18	DMI	ARPEGE	HIRHAM	green
19	OURANOS	CGCM3	CRCM	purple
20	GKSS	IPSL	CLM	pink

Table 4.1: List of the models evaluated in this study. Note that we do not differentiate between HadCM3 different sensitivities (Q0, Q3 and Q16) and consider them as the same model. *Model evaluated but excluded from the analyses and from the Figures (see Section 4.3.2).

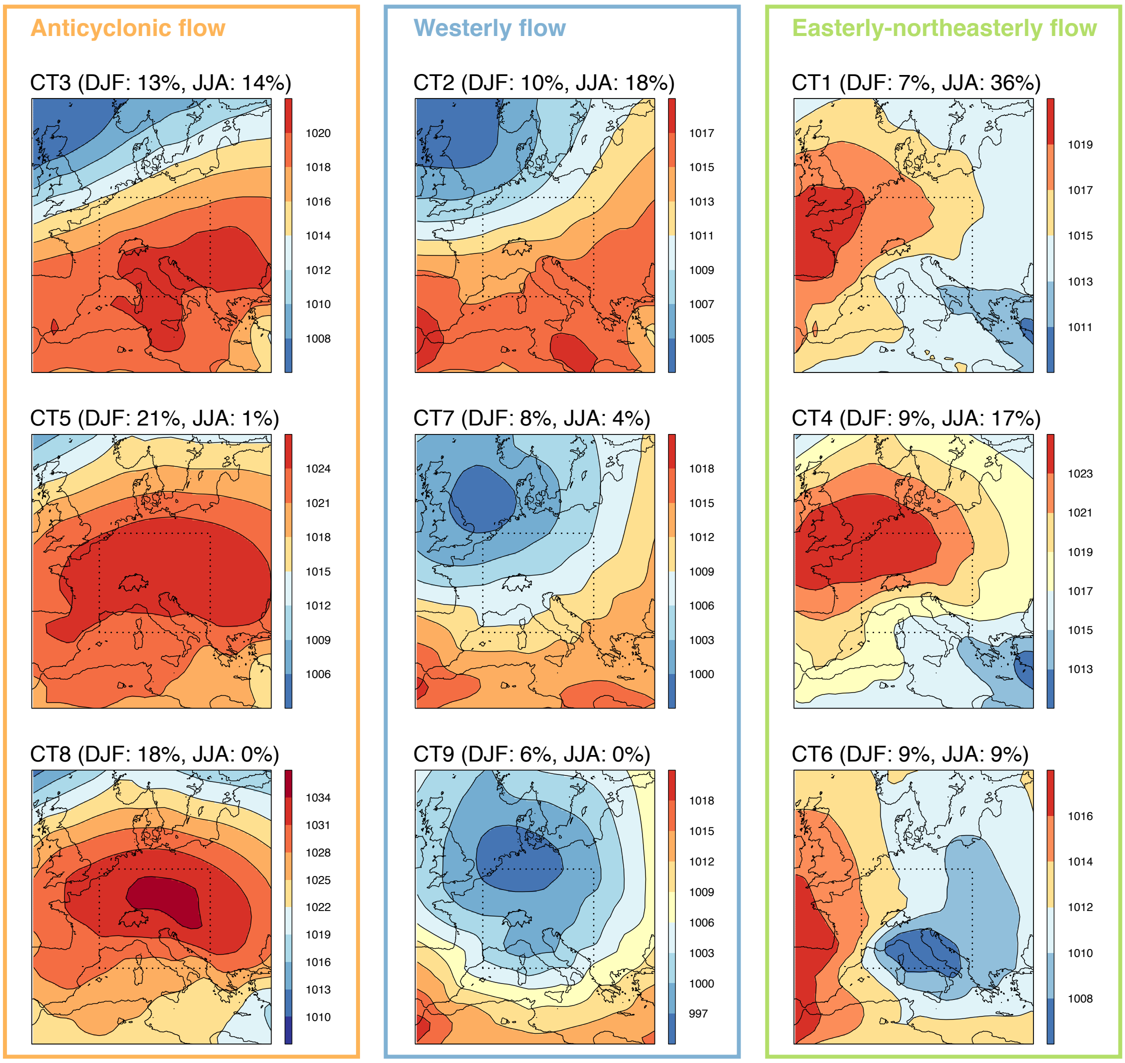

Figure 4.1: Mean sea level pressure [hPa] for the nine centroids of the PCACA09 classification determined over the Alps domain (dotted rectangle) including Switzerland (closed continuous line within the rectangle). The frequency of occurrence of each CT in DJF and JJA is indicated in brackets. In the text and figures we refer to the CTs using their dominant flow and the following colors: anticyclonic (orange), westerly (blue) and easterly-northeasterly (green).

which to explore how well models capture the combined influence of the Alpine range and synoptic circulation on atmospheric conditions. Further, the size of the domain represents a compromise between an evaluation of specific grid cells, which may lead to an over-interpretation of model outputs because this resolution is finer than the model effective resolution, and an evaluation over a larger domain, in which opposite biases in different locations are more likely to partially mask each other.

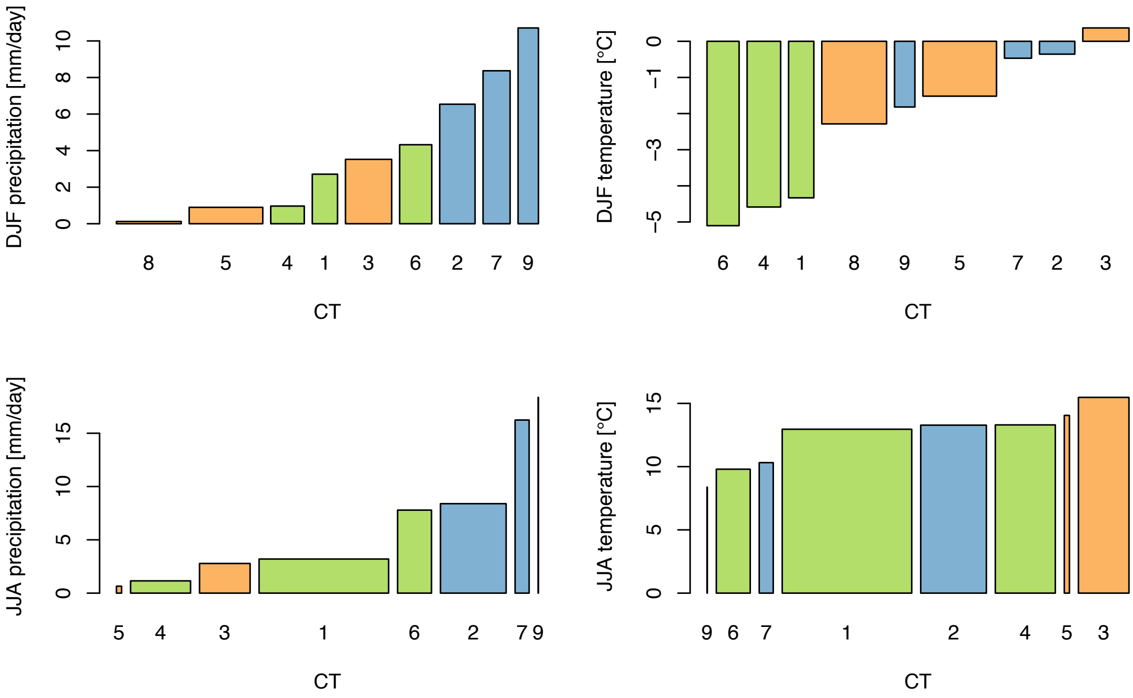

Figure 4.2: Mean daily precipitation amount (left) and temperature (right) observed over Switzerland in winter (top) and summer (bottom) for each CT (determined using ERA40 sea level pressure). The color of the bars corresponds to the CT clusters (see Figure 4.1), their width is proportional to the frequency of occurrence of each CT, and they are ranked on the basis of their mean precipitation or temperature. Note that no days are attributed to CT 8 in JJA.

4.2.2 Characterization of the synoptic circulation

For each GCM-RCM combination, we characterized the daily synoptic situations using an objective CT classification based on a combination of principal components analysis and hierarchical cluster analysis (PCACA, Philipp et al., 2010, 2014). The classification relies on mean sea level pressure (MSLP) fields within an area representative of the Alps (domain D06 in Philipp et al., 2010, see also Supporting Figure 4.1) and exists in three variants, using 9, 18 and 27 CTs. Schiemann and Frei (2010) compared 71 classification schemes and found that the 9 CT variant of the PCACA scheme captures winter daily precipitation over the Alps in a more reliable way than most other classification schemes with 9 CTs and also better than a number of classification schemes using 18 or even 27 CTs, although its performance is weaker in summer than in winter. In this study we utilized the 9 CT variant of the PCACA scheme. CT classes can be determined at the seasonal level season, but here we chose the classes determined on the annual level. The centroid of each CT and the time series of ‘observed’ CTs were derived from ERA40 reanalysis (Uppala et al., 2005). Using another reanalysis data set might have led to different CT time series, but we assume these differences to be overall smaller than the differences between the reanalysis and the climate simulations (Riediger and Gratzki, 2014). We then assigned the simulated conditions of every day to the nearest CT in terms of Euclidian distance to its centroid (Rohrer, 2013). This means that we used the same centroids (determined using ERA40) to classify and then compare all the GCM-RCM simulations. The investigated period was 1980-2001, with the end being determined by the availability of the ERA40 data set. The initial analysis showed that the model combination ECHAM5-HIRHAM presents unreasonably large biases over Switzerland in winter and hence this GCM-RCM was excluded from further analyses (more details in Section 4.3.2).

4.2.3 Bias propagation from the synoptic to regional scale

We hypothesized that part of the biases in precipitation and temperature comes from circulation biases (van Ulden et al., 2007), in particular from the under- and over-representation of specific CTs. For instance, a cold bias in the regional simulations may originate from an overly frequent occurrence of CTs associated with cold conditions. We considered both the origin of the atmospheric flow and average temperature and precipitation over Switzerland. This enabled us to determine which of the 9 CTs lead to cold or warm, wet or dry conditions (Figures 4.1 and 4.2). As can be expected, in both winter and summer the CTs leading to the highest average daily precipitation correspond to westerly flows, advecting moist air from the Atlantic Ocean. In summer high pressure systems over, or close to, Switzerland lead to the warmest conditions, while in winter the coldest conditions are caused by continental easterly-northeasterly flows. Given the good agreement between the large-scale flow and the mean precipitation and temperature, we decided to combine the 9 CTs into 3 types of dominant flow: westerly, easterly-northeasterly and anticyclonic flow (Figure 4.1). This choice has a double motivation. Firstly it was important for visualization and interpretation purposes to reduce the dimensionality of the problem while retaining the main CT differences, and secondly, since we perform CT-dependent bias-adjustment, we needed enough data for each CT, which was not guaranteed for individual CTs with a low occurrence rate.

To assess how much bias can be explained by biases in CT frequency, we used two data sets: ERA40-driven RCM simulations and time series constructed by resampling GCM-RCM simulations. Using a reanalysis to force an RCM improves fluxes at the domain boundaries, which leads in particular to a better representation of the synoptic situation. To isolate the influence of biased CT frequency on mean precipitation and temperature in RCM simulations, we resampled with replacement from the GCM-RCM runs. For each CT, we computed how many days ($N_{CT,OBS}$) were dominated by this CT in the observations and extracted $N_{CT,OBS}$ days from the $N_{CT,SIM}$ days dominated by the same CT in the GCM-RCM simulation. We then added the days from the other CTs following the same procedure. This leads to inconsistencies in the temporal sequence of the time series, but this is not an issue in our case since we here restrict our attention to mean values.

4.2.4 Bias-adjustment using quantile mapping

We bias-corrected the RCM simulations by applying quantile mapping using two setups. For both setups quantile mapping was implemented using empirical quantiles (Gudmundsson et al., 2012) and was applied on daily precipitation and temperature averaged over Switzerland. In the first 'standard' setup, the quantile mapping adjustment is conditional on the season, so two adjustments were used, one for summer, one for winter. In the second 'CT-dependent' setup, the adjustment is conditional on both the season and the CT, so six adjustments were used for temperature and six for precipitation. To calibrate the CT-dependent quantile mapping, we used all days with a particular flow during a particular sesaon (e.g., anticyclonic flow in DJF) from both the observations and a particular GCM-RCM run. The differences between the empirical cumulative distribution functions for these two data sets were then used to establish the transfer functions on which the quantile mapping method relies. This idea to apply a CT-depend bias-adjustment is inspired by B·rdossy and Pegram (2011), who observed that different CTs can lead to different biases in RCM simulations and, based on the fact that different CTs correspond to different atmospheric processes, proposed that the bias-adjustment method should be CT-dependent.

4.3 Results

4.3.1 Circulation biases and within circulation type biases

Different CTs correspond to different regimes in terms of precipitation and temperature, and overall the models capture these differences reasonably well (Figure 4.3). For instance, when in winter continental air is advected along the Alps by an easterly-northeasterly flow, the average temperature is lower than if the region was under the influence of a westerly or anticyclonic flow (Figure 4.3b). This is captured by the models and they also capture differences in CT frequency, for instance that in winter, days with a dominant easterly-northeasterly flow are less frequent than those with a dominant anticyclonic regime (Figure 4.3b). Yet there is some considerable spread of the simulations among the observations, which reflects the existence of biases in the frequency (circulation biases) and the precipitation intensity/temperature of each CT (within CT biases). The contribution of these two kinds of biases to the mean seasonal bias is variable. For instance, in winter, climate models tend to overestimate both the frequency of westerly situations and their mean precipitation intensity (Figure 4.3a). Since this circulation type is associated to high precipitation intensities (Figure 4.2), a small bias in the frequency leads to an important bias in the seasonal average (see the steep isolines in the upper part of Figure 4.3a). In contrast, an overestimation of the frequency of days with drizzle leads to much smaller changes in the seasonal average (see the flatter isolines in the lower part of Figure 4.3a). Further, as pointed out by Bárdossy and Pegram (2011), different CTs may show different biases. For instance in summer, models tend to overestimate precipitation amount on easterly-northeasterly flows, but typically underestimate it for westerly situations (Figure 4.3c).

Winter precipitation in Switzerland is overestimated by 18 out of 19 models (black symbols in Figure 4.3a), a systematic bias also reported by Fischer et al. (2012), and leading to the simulation of an unrealistically large snow accumulation and overall water resources. The overestimation of winter precipitation by the models is too large to be explained by precipitation undercatch alone. Further, it is unlikely that independent models show such a consistent positive bias. Instead, we hypothesize that this bias stems from the synoptic fields inherited from the GCMs (van Ulden and van Oldenborgh, 2006), with 2 GCMs driving as many as 13 RCMs, i.e., from a lack of independence of the model runs.

4.3.2 RCMs' response to circulation biases

To explore the influence of the synoptic field on RCM simulations, we represented the bias in temperature and precipitation as a function of the frequency bias in one CT (Figure 4.4). Two features are particularly clear. Firstly, the clustering of points of the same colors indicates that the driving GCM has a considerable influence on the circulation bias, both in winter and summer. Although RCMs respond to the biases differently, we find that the CTs they simulate are to a large extent prescribed by the driving GCM. GCMs explain the majority of the variance of the bias in the CT frequency, both in winter and summer (from 57% for the north-northeasterly flow in winter to 88% for the anticyclonic flow in summer,

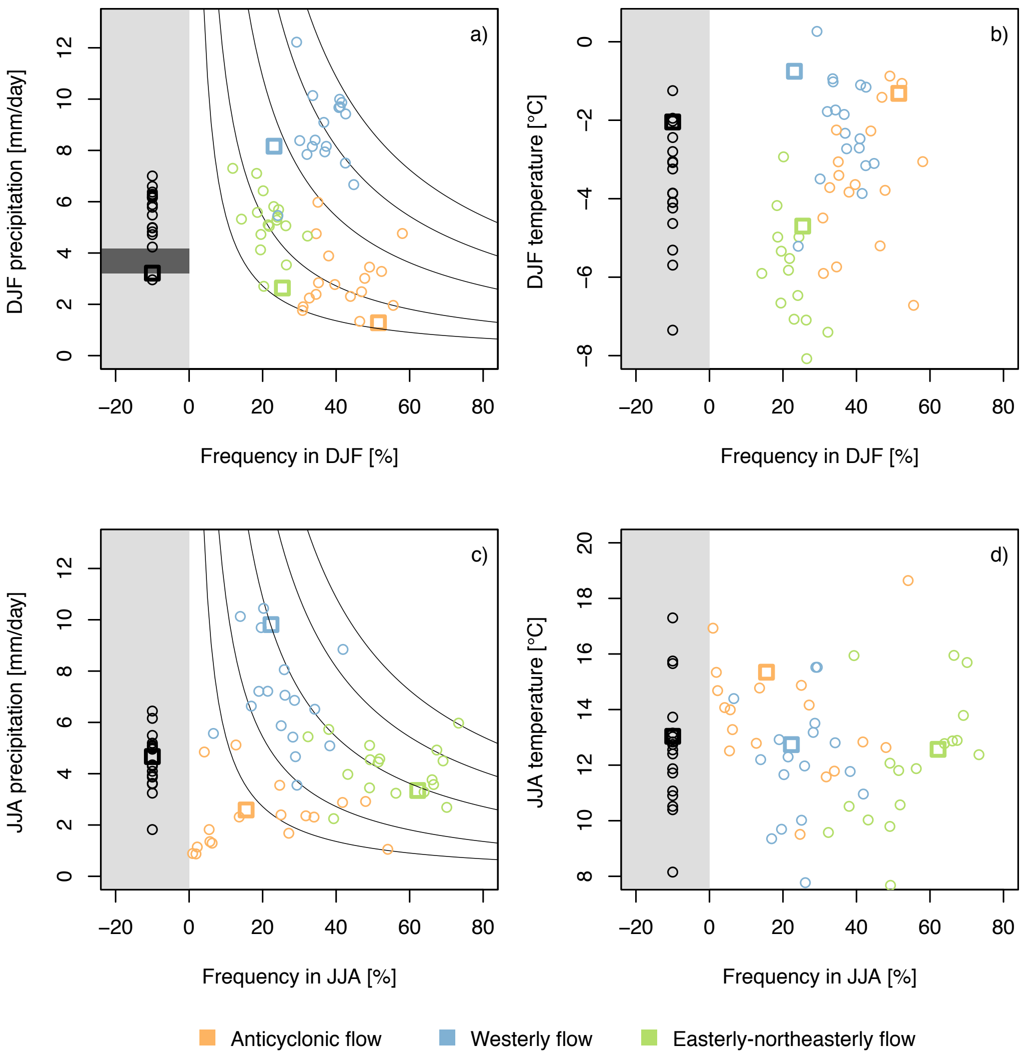

Figure 4.3: CT frequency, temperature (left) and precipitation (right) in winter (top) and summer (bottom). The squares represent observations, the open circles simulations from the 19 GCM-RCMs. Black symbols correspond to seasonal averages and the colored ones to CT averages. The isolines indicate the combinations of the mean precipitation and frequency leading to the indicated seasonal (3-month) precipitation amounts. The dark gray area in panel a) represents the estimated undercatch of winter precipitation (30%).

not shown). Secondly, in winter, there is an almost systematic overestimation of the precipitation and of the frequency of westerlies (Figure 4.4a), and the models overestimate the precipitation by 71% and the frequency of westerlies by 57% on average. The significant correlation between these two variables (r = 0.57) is consistent with the fact that an overestimation of westerlies leads to too much moist air being transported from the Atlantic and precipitating in the Alpine region. A similar conclusion was drawn by (van Ulden et al., 2007) for the previous generation of the RCM runs for Europe (PRUDENCE project, Christensen and Christensen, 2007). Their study focused on Central Europe, for which they reported an overestimation of the west component of the geostrophic wind in winter, and established its contribution to wetter winters. Although here we did not consider the

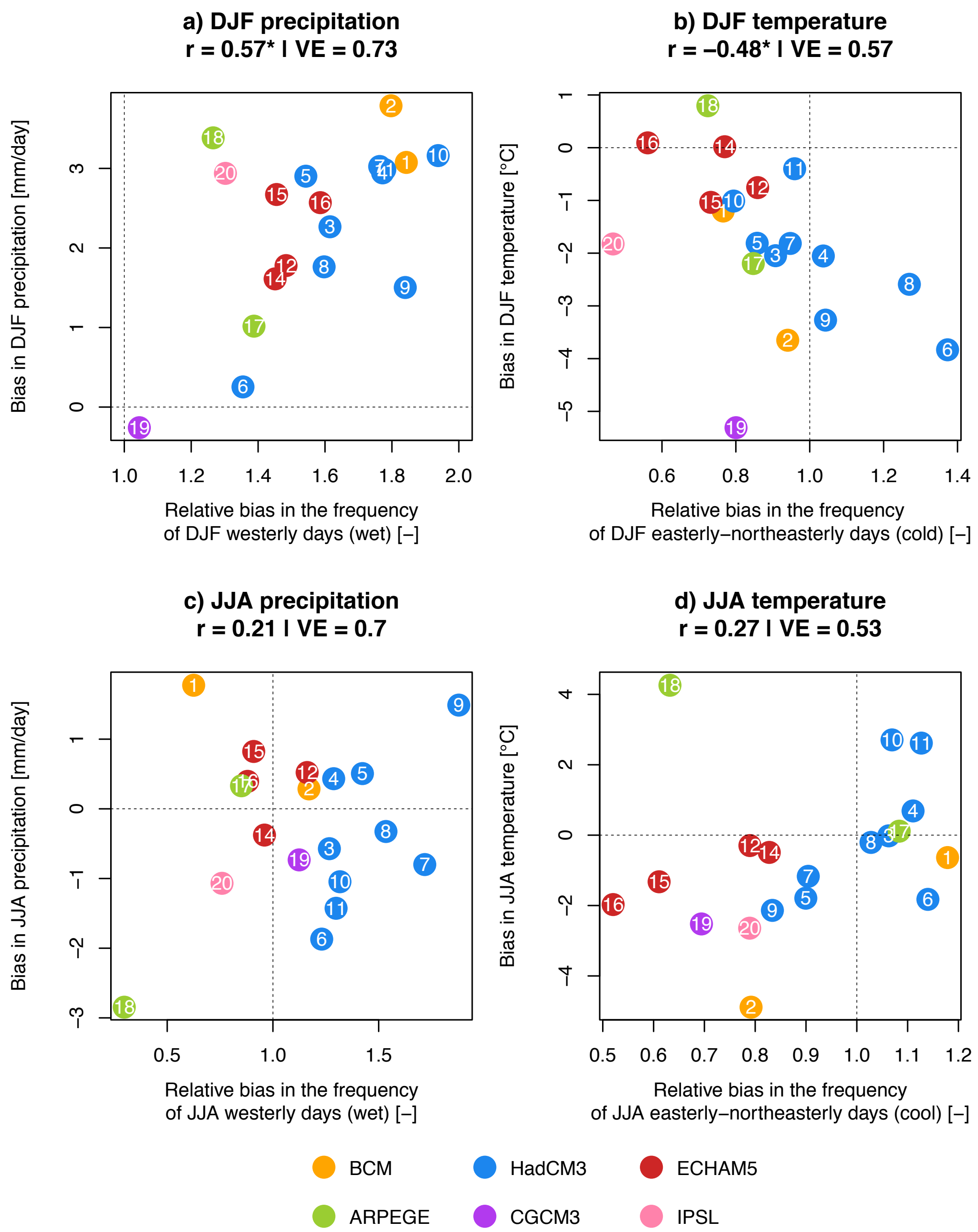

Figure 4.4: Bias in the frequency of a selected flow situation and in mean seasonal precipitation and temperature for the 19 GCM-RCMs. The Pearson's product moment correlation coefficient r (star indicates significance at the 5% level) and the explained variation EV of the frequency bias by the GCMs are indicated. Dots of the same color represent RCMs driven by the same GCM, the numbers refer to the different GCM-RCM combinations (Table 4.1).

strength of the flow, their results are compatible with the overestimation of the frequency of the westerlies we report, which suggests a persistent bias. Note that that precipitation is overestimated for all three flow situations (Figure 4.3a), which indicates general issues with the representation of winter precipitation by the models in our study area.

As mentioned before, we excluded the model ECHAM5-HIRHAM from this study. We found that the winter precipitation it simulates over Switzerland is about 328% of the observed amount and that 71% of its winter days are westerlies, which should be compared to the observed frequency of 23%. This agrees with the connection of biases in CT frequencies and precipitation amounts, but because of these particularly large biases, we decided to exclude this model from all the analyses and figures presented in this study. The tendency of ECHAM5-HIRHAM to overestimate winter precipitation to a larger extent than the other ENSEMBLES model chains was also reported for Spain by Turco et al. (2013), but they did not investigate the reasons behind this bias.

The bias in winter temperature is significantly negatively correlated ($r = -0.48$) with the bias in the frequency of easterly-north easterly flow that, in the models and in the observations, lead to below average temperature in Switzerland (Figure 4.4b). In other words, these two biases are consistent. The main axis of the cloud of points in Figure 4.4b does not include the origin, which is consistent with the underestimation of winter temperature for almost all the model-CT combinations Figure 4.4b.

The influence of synoptic situation is much weaker in summer, the correlations between the circulation biases and the surface variables (mean precipitation and temperature) are not significant. RCMs driven by the same GCM show similar circulation biases, but these can lead to quite different temperature biases. This likely reflects that processes at mesoscale to the regional spatial scale, such as soil moisture feedbacks and convective precipitation, are more influential during the summer months. Figure 4.4d also highlights that some models provide reasonable estimates of mean summer temperature, despite boundaries conditions with clear deficiencies. For instance, the frequency of situations dominated by an anticyclonic circulation is clearly underestimated by HadCM3 and overestimated by ECHAM5 (blue versus red dots in Figure 4.4d). Yet some RCMs driven by these GCMs manage to capture the mean observed temperature surprisingly well. This raises the question whether these models get the right answer for the right reasons, and suggests that projected changes in CTs should be interpreted with caution, given the important biases under present climate conditions.

The results in Figure 4.4 depict the correspondence between biases in mean precipitation and temperature, and a single CT. To account for the three CTs at once, and to estimate the potential benefits of reducing CT frequency biases, we used constructed time series based on the resampling of GCM-RCM simulations. Figure 4.5a shows that the correction of the CT frequency leads to a decrease of winter precipitation bias by 32% (mean across the models). This provides further support to the idea that biases in CT frequency directly contribute to the wet winter bias in Switzerland. The overestimation of winter precipitation is reduced even further when ERA40 is used to force the RCMs. Note however that in contrast to winter precipitation, winter temperature barely benefits from a correct CT frequency (Figure 4.5b). The reason is that even if precipitation is overestimated on westerly days and anticyclonic days, removing westerly days (mean daily precipitation ∼8mm) and replacing them by anticyclonic days (mean daily precipitation ∼2mm) will lead to a decrease of the seasonal precipitation amount. This is also illustrated in Figure 4.3a): the isolines are steeper around the blue points than around the orange points, meaning that changing the frequency of westerly days has a larger impact on seasonal precipitation amount than changing the frequency of anticyclonic days by the same amount. In summer there is almost no benefit from correcting CT frequency biases (Figure 4.5c and 4.5d).

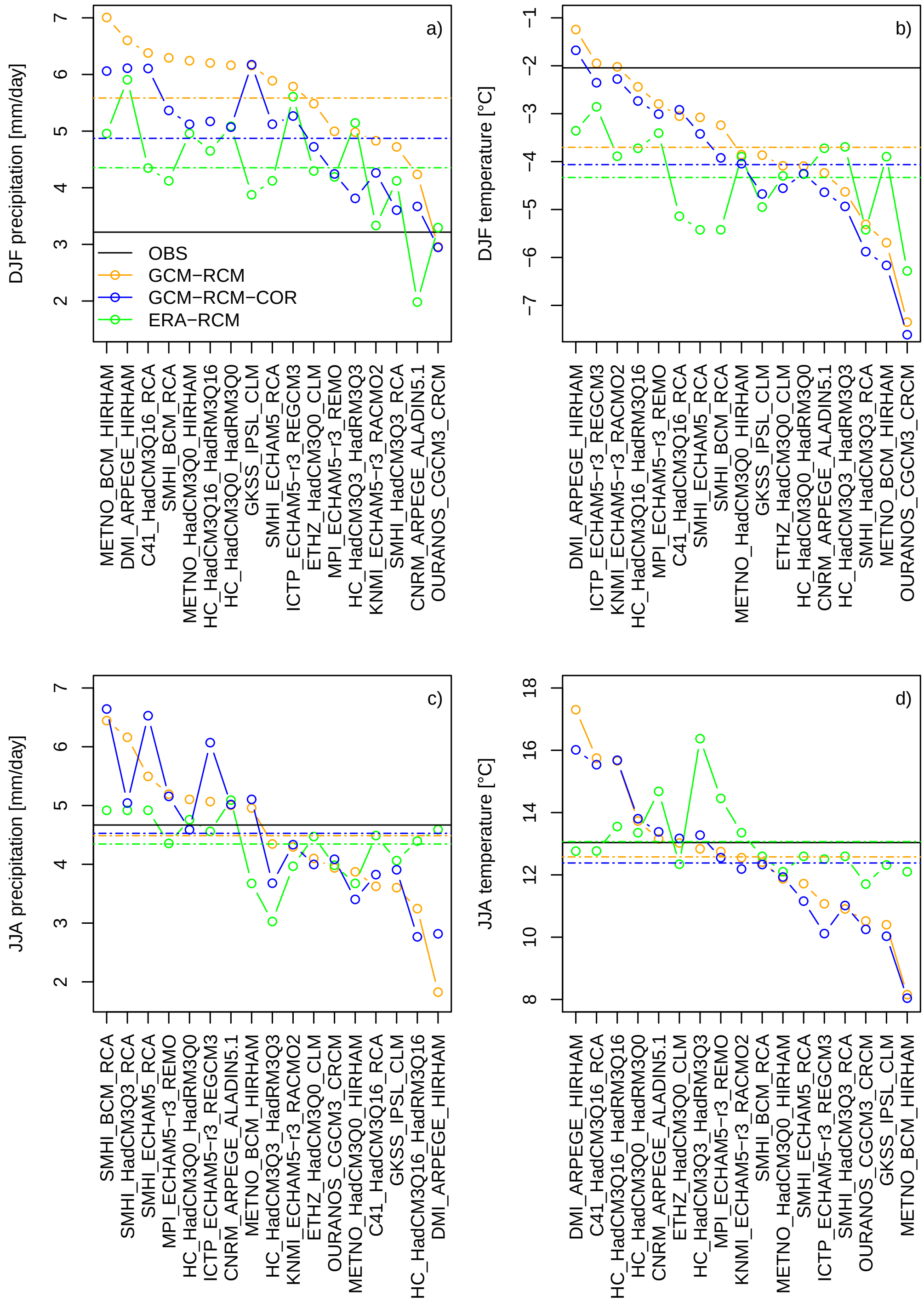

Figure 4.5: Assessment of the potential benefit of correct CT frequencies. The blue line indicates RCM simulations in which only the CT frequency is corrected (by resampling), while the green line indicates RCM model simulations in which boundary conditions (in particular SLP, on which the CT classification is based) are prescribed by ERA40. The dashed colored lines indicate the seasonal mean across the models. Models are sorted based on the mean seasonal precipitation (left column) or temperature (right column) of the raw GCM-RCM simulations (orange line), the GCM-RCM combination being indicated on the x-axis.

4.3.3 Residual errors in bias-corrected time series

We tested the sensitivity of bias-corrected time series by applying quantile mapping using two setups: one conditional on the season, and one on both the season and the CT. Both variants perform well at the tasks they were trained for, but their lack of consideration for the physical processes leading to biases prevents them from correcting for other aspects. As expected, when quantile mapping is calibrated for each season, the mean seasonal precipitation is correct (Figures 4.6a to 4.6c), and when it is calibrated for each CT, the precipitation on each of the CT is correct (Figures 4.6e to 4.6f). However, significant biases remain in the corrected time series of winter precipitation, as quantile mapping introduces compensation errors. For instance, for the bias-corrected simulations to match the observed seasonal precipitation despite the frequency overestimation of westerly flow, standard quantile mapping biases low the precipitation on westerly flow situations (ε_2 in Figures 4.6e to 4.6f). Similarly, the seasonal precipitation is overestimated (ε_1 in Figures 4.6a and 4.6c) when a CT-dependent correction is applied because the overestimation of westerly days is not accounted for.

These residuals errors are present for all models. Figure 4.6g shows that ε_1 and ε_2 tend to increase in absolute terms with circulation biases, represented here by the bias in the frequency in the winter westerly flows. For instance, the model HadCM3-RCA3 (#10) presents both the largest residual error in the seasonal winter amount (43%) and the largest overestimation of the westerly frequency (93%) of all models considered. Note the particular situation of the models IPSL-CLM (#20) and HadCM3-RRCM (#6). Their circulation bias in winter is lower than the average, but their residual error ε_2 is larger than that of the other models. This apparently results from the misdistribution of precipitation among three circulation types in the raw (uncorrected) model runs. In the observations, the ratio of mean precipitation on a westerly day to a easterly-north easterly day is 3.1:1, but it is largely underestimated in both IPSL-CLM (1.1:1) and in HadCM3-RRCM (0.8:1). The correct ratio is reached when CT-based correction is applied, but the standard quantile mapping fails to reestablish the balance between the different regimes. Even after the correction it still rains more in average on easterly-north easterly days than on westerly days in HadCM3-RRCM.

As already stressed, bias-adjusting DJF precipitation using standard quantile mapping leads to an degradation of the simulations of precipitation on westerly days (Figure 4.7a). But in other cases, the situation on specific CTs is improved by the bias-adjustment of the overall seasonal conditions. This applies for instance to DJF temperature on easterly-northeasterly days (Figure 4.7b) and to JJA temperature on anticyclonic days (Figure 4.7d). Note that JJA precipitation on westerly days is improved by the correction of JJA precipitation for some models, but it is degraded for other models (Figure 4.7d). Overall, the four panels of Figure 4.7 illustrate that while quantile mapping leads by construction to a good match between the distribution of the observations and that of the post-processed time series, it also alters other aspects of the model simulations. Whether these alterations are beneficial or detrimental depends on variable of interest and the GCM-RCM.

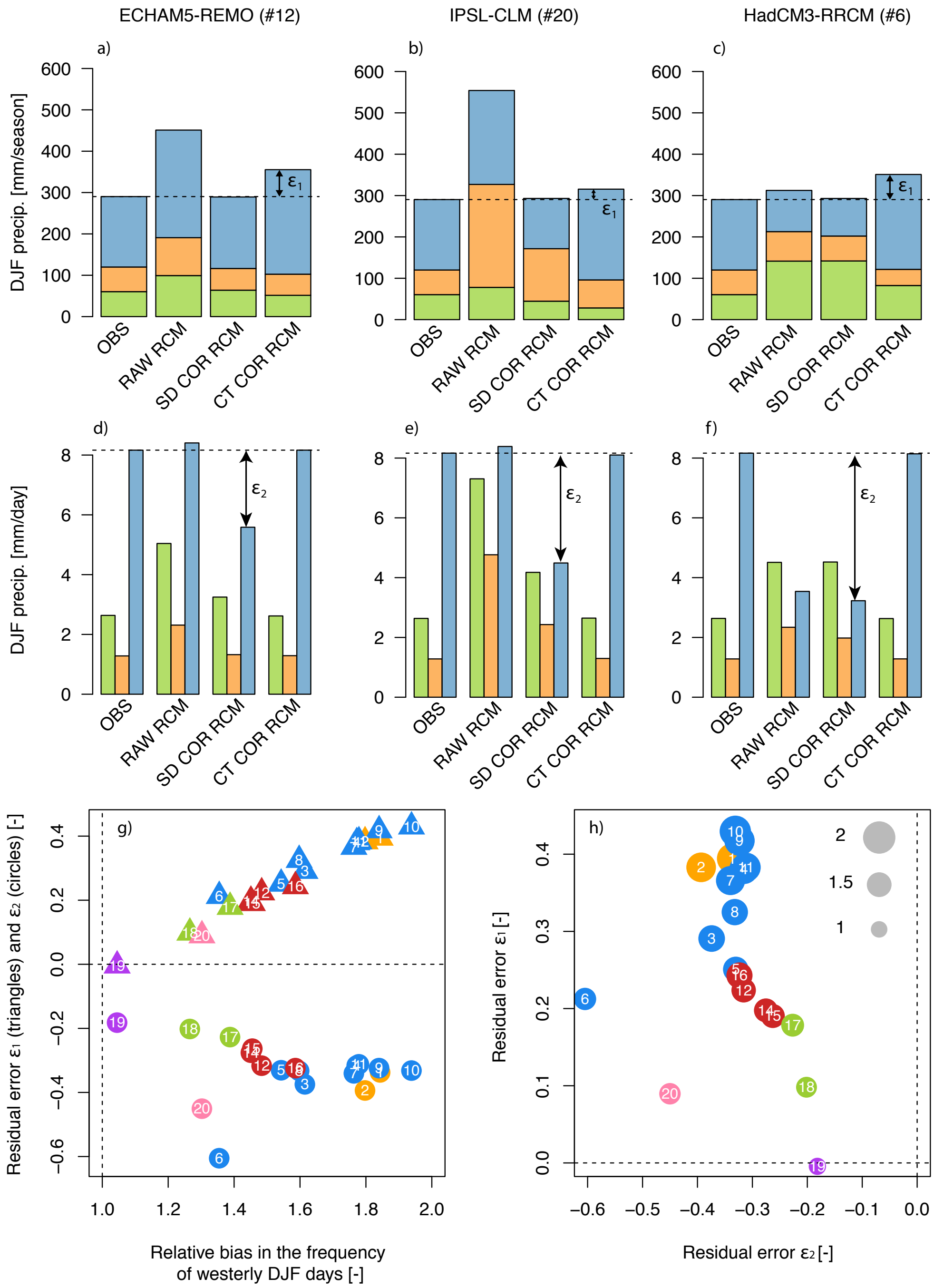

Figure 4.6: An illustration of residual errors in bias-corrected time series and of their relation with large-scale circulation biases. a) to f) Winter precipitation as observed (OBS), as simulated by three models without bias-adjustment (RAW RCM) and after bias-adjustment using standard and CT-dependent quantile mapping (SD COR RCM and CT COR RCM, respectively); the colors correspond to the dominant flow (see Figure 4.1). g) Residual error in mean winter precipitation after CT-dependent bias-adjustment (ε_1) and in mean winter precipitation on westerly days after standard bias-adjustment (ε_2), expressed relative to the observed value; the colors and numbers correspond to the GCMs and RCMs (see Table 4.1). h) Same information as in g), but with more emphasis on the driving GCM; the size of the dots is proportional to the bias in the frequency of westerly flow in winter.

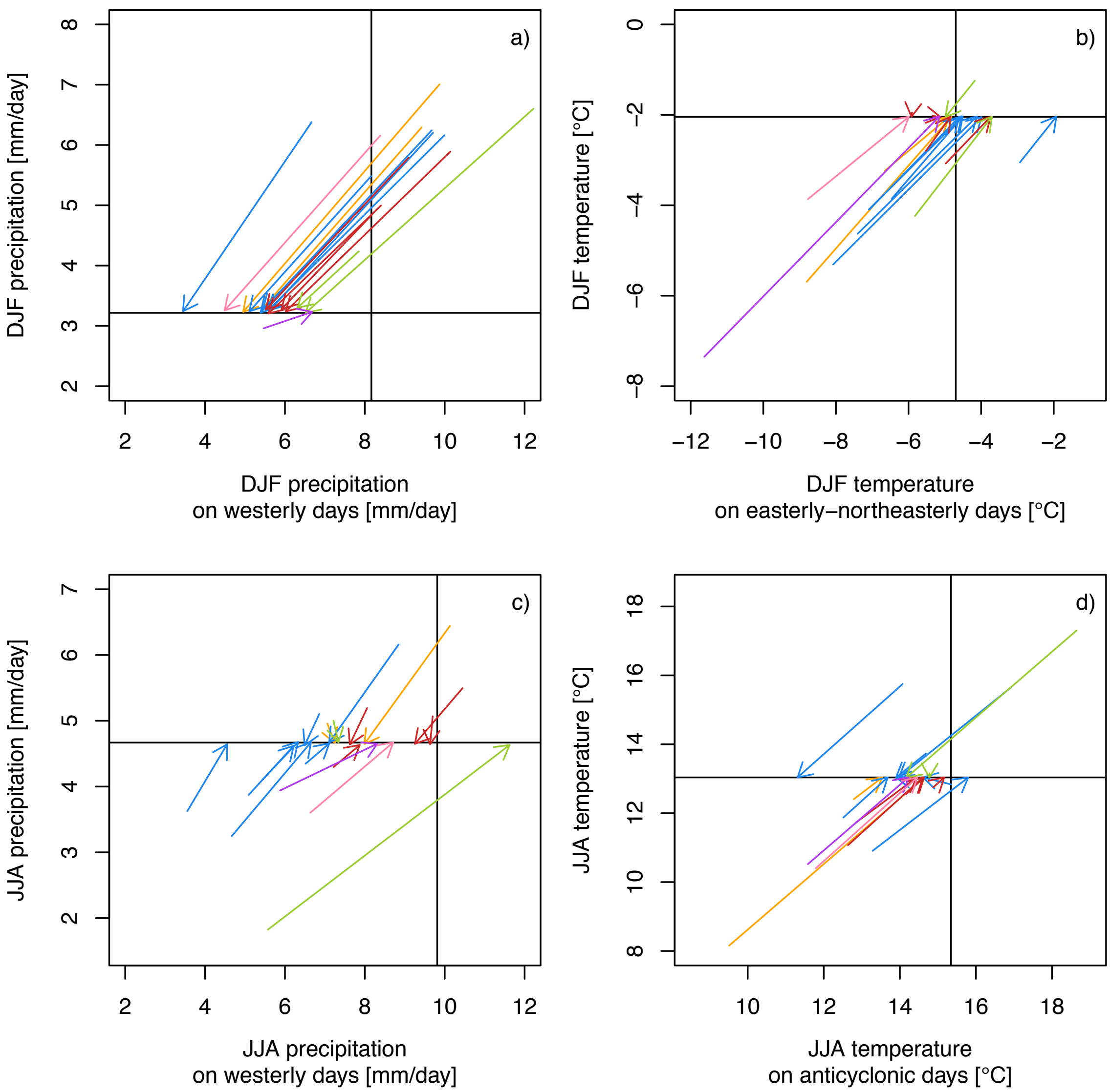

Figure 4.7: Residual errors for selected CTs after standard quantile mapping correction. The y-axis indicates the variable quantile mapping has been trained to correct at the seasonal scale, while the x-axis represents the same variable for a specific CT, that has not been considered in the calibration of the quantile mapping method. The solid black lines represent the observed value of the two variables.

4.4 Discussion

4.4.1 Implications for bias-adjustment and impact modeling

Overall, quantile mapping can enable substantial bias reductions. Yet we find that when circulation biases are present, they lead to significant residual errors in the time series. In such situations the limitations of ad hoc adjustments not accounting for processes clearly appear. RCMs inherit circulation biases from their driving GCM, which leads to biases in surface variables crucial for impact modeling. These circulation biases illustrate the difficulty to capture the present and future climate dynamics using models (Shepherd, 2014). They also represent a challenge for the interpretation of climate projections, and for the bias-adjustment of RCM simulations. We showed that when circulation biases are large, they lead to significant residual errors in bias-corrected time series. Yet, although we were able to identify their contribution to biases in winter precipitation and temperature, it is unclear how to correct them by post-processing, since they are deeply rooted in the simulations. Indeed, as atmospheric circulation is at the core of climate projections, it needs to be handled with care, and pragmatic approaches often used when preparing climate models outputs for impact simulations might not be appropriate. It remains to determine how best account for circulation biases, in particular how we could correct them a posteriori without compromising the reliability of the simulations. Concretely, how should days of overrepresented CTs be selected and removed, and which days of underrepresented CTs should be selected to replace them?

One way to progress would be to embrace the fact that different models have different deficiencies. We argue that model errors should be better diagnosed (for instance using CTs) and this should guide the choice the bias-adjustment method. That is, instead of using the same bias-adjustment for all the models of the ensemble in a 'one fits all' approach, model-specific bias correction methods could be used. For instance, choosing between a standard or a CT-based bias-adjustment method may be a choice that is model dependent: in the case of HadRCM3-RRCM, the biases within CTs are so important (Figure 4.6f), that a CT-based approach may be considered to give better result than the standard approach.

When it comes to impact modeling, a key question is how much residual errors in bias-corrected time series matter. This depends on the sensitivity of the impact model to the bias-adjustment (Muerth et al., 2013; Teng et al., 2015) and on the other sources of uncertainty involved (Chen et al., 2011; Addor et al., 2014). In some cases, the time-series might be deemed 'good enough', for instance if their residual errors are smaller those in the observations, resulting from interpolation and measurement processes. In others, for instance when the impact model is highly distributed or if the focus is on a few extreme events, then these residual errors may become a concern.

Applying quantile mapping in its standard form leads to an overestimation of the mean precipitation falling on days dominated by an easterly-northeaserly flow, which are in average more than 2°C colder than days dominated by the two other flows. In other words, although mean seasonal precipitation is correct, too much of it falls on cold days. This will lead to an overestimation of the snow pack. This aspect is of clear importance for hydrological modeling in Swiss catchments.

4.4.2 Implications for model development, interpretation and evaluation

RCMs are designed to disaggregate GCM simulations in space and time. They however do not have the purpose, nor the ability as illustrated by this and other studies, to correct biases in the frequency of synoptic situations. It implies that even if the resolution of the RCMs used in this study was further increased, the frequency of westerlies simulated by the GCM (and hence the water vapor introduced at the domain boundaries) would remain overestimated, and consequently winters would remain too wet. This is a case in which the added value of dynamical downscaling is limited by the boundary conditions provided by the GCM (Racherla et al., 2012), and RCM simulations reflect more GCM biases than regional features. The fact that GCMs can introduce significant biases in RCM simulations underscores the importance to not only evaluate RCM simulations using a reanalysis-RCM setting (i.e., with close to correct synoptic fields), but also using a GCM-RCM setting (i.e., with potentially significantly biased synoptic fields). The former allows for the separation between downscaling deficiencies and deficiencies of the boundary forcing and is a crucial step in model development and tuning. Yet since RCMs are overwhelmingly used to gain insight into future conditions, for which no re-analysis is available, it is essential to also evaluate them in a GCM-RCM setting, to test how well they cope with potentially significant circulation biases. We argue that this is particularly true in the framework of bias-adjustment studies and impact assessments.

Given the substantial biases that can be introduced by GCMs, we advocate for RCM evaluations to be carried out for RCMs driven by a GCM (i.e., with potentially significantly biased synoptic fields) rather than by a reanalysis (i.e., with close to correct synoptic fields). We recognize the value of reanalysis-driven RCM simulations as they enable to concentrate on the RCM. But since RCMs are overwhelmingly applied under future conditions, for which no re-analysis is available, we argue that their skill for climate studies is most relevant when it is evaluated in a GCM setting (Kerkhoff et al., 2014). In other words, reanalysis-driven RCMs do not allow us to test how well RCMs cope with significant circulation biases. We propose that for the evaluations of downscaling and bias-adjustment methods, the GCM setting should be favored, as it is key to better assess how well these methods handle circulation biases.

Finally, the importance of the circulation biases under present conditions for the models investigated here implies that caution should used when be interpreting projected changes in the frequency of circulation types. Now that the new CORDEX generation of GCM-RCM simulations starts being used for regional assessments of climate change and impact studies, we recommend to first assessing the reliability with which CT are simulated in these new simulations, before inferring impacts from projected changes in their frequency. The GCMs involved in this study are part of the climate model generation CMIP3. The comparison performed by Perez et al. (2014) suggests that models of the following generation (CMIP5) capture better the frequency of weather types over the north-east Atlantic region. However, it remains to be evaluated whether this improvement is significant enough to lead to a decrease of biases in regional precipitation and temperature.

4.5 Conclusions and outlook

Accounting for the processes leading to biases is essential to progress with bias-adjustment, but these processes are difficult to pinpoint. One way to identify them is to perform a CT-dependent evaluation of climate models. This enables to relate the synoptic scale to the regional scale, and can reveal model deficiencies overlooked when performing model evaluation at the seasonal time scale. We used a CT-dependent evaluation to explore precipitation and temperature over Switzerland, which revealed that the overestimation of the frequency of westerly situations was a significant contributor to the overly wet conditions simulated by the large majority of the ENSEMBLES GCM-RCMs. However we could not explain biases in other variables (e.g., summer temperature), which require a focus on other scales and variables (e.g., soil moisture, Bellprat et al., 2013). Further, our study investigated mean precipitation and temperature over Switzerland, but future studies could concentrate on smaller, well-chosen areas to address questions such as how well precipitation on different sides of the Alps is reproduced under different synoptic situations or how well wet or dry extremes in relation to different CTs are captured by the GCM-RCMs.

As for bias-adjustment, we propose that a better understanding of what we correct is necessary. It is well-established that bias-adjustment only improves selected aspects of climate simulations and leave inconsistences in the post-processed time series, which we refer to as residual errors. However, we go beyond this simple fact, and illustrate how the existence and the amplitude of these residual errors can be related to misrepresentation of specific atmospheric processes (i.e., the frequency of specific CTs). We show that compromises will be necessary, since the correction of one aspect of the RCM simulation can lead to a degradation of other aspects.

We used circulation types (CTs) as a diagnostic tool to establish links between the synoptic and regional scale, and thereby to gain insights into the origin of biases in RCM runs over Switzerland. GCMs introduce significant biases in the frequency of CTs simulated by RCMs, which lead to biases in simulations of winter precipitation and temperature. We show that winter precipitation is particularly overestimated, which is a concern for hydrological modeling in snow-dominated basins. The consistence of this bias across the 20 GCM-RCM chains evaluated here comes mainly from the overestimation of the westerly frequency common to the five driving GCMs. Our results suggest that a better simulation of the synoptic circulation in winter could contribute to reduce this wet bias. Significant circulation biases also exist in summer, but we do not find them to be correlated to bias in precipitation and temperature, implying that those regional biases originate from the misrepresentation of physical processes at smaller spatial scales.

We propose that the choice of the bias correction method should depend on two main criteria: i) the origin of the biases, which we argue must be better diagnosed and ii) the purpose of the bias-corrected time series. In the context of precipitation over Switzerland, it is important to determine what is most valuable: the right seasonal precipitation, or how it is distributed among different synoptic circulations? The answer to this question is application dependent and will rely on the experience of impact modelers.

Overall, circulation biases represent a real challenge for both dynamical downscaling and the bias-adjustment of RCM simulations, because they are deeply rooted in the atmospheric circulation simulated by GCMs. As more recent GCMs improve the representation of atmospheric circulation, this should contribute to the reduction of RCM biases. Further research is necessary to better understand the processes behind biases in GCM-RCM simulations, and to clarify how downscaling and bias-adjustment methods should account for them.

Acknowledgements

We are grateful to the contributors to the COST733 Action ‘Harmonisation and Applications of Weather Types Classifications for European Regions’ for their systematic analysis of CT classifications and for making the classification software cost733class available (Philipp et al., 2014). We thank the EU FP6 Integrated Project ENSEMBLES projects for the GCM-RCM simulations (http://ensemblesrt3.dmi.dk/) and the Swiss Federal Office of Meteorology and Climatology, MeteoSwiss for the RhiresD and TabsD data sets (these data sets can be requested from MeteoSwiss). The CT classification of the GCM-RCM simulations is available from Marco Rohrer (marco.rohrer@giub.unibe.ch). This article benefited from constructive comments from three anonymous Reviewers and from stimulating and inspiring discussions with Renate Wilcke and Andreas Prein. NA thanks the University of Oregon, USA for hosting him while finishing this study. This work was supported by the Swiss National Science Foundation (grant 200021_131995) and by the University of Zurich (University Research Priority Programs on Global Change and Biodiversity).

Chapter 5

Bias correction for hydrological impact studies - beyond the daily perspective

Nans Addor[1], Jan Seibert[1,2]

[1] Department of Geography, University of Zurich, Zurich, Switzerland
[2] Department of Earth Sciences, Uppsala University, Uppsala, Sweden

Published in Hydrological Processes, 2014
doi:10.1002/hyp.10238

5.1 Emergence of bias correction methods in impact studies

Hydrological simulations driven by climate projections are a central tool to assess the availability of future water resources and changes in droughts and floods (Jiménez Cisneros et al., 2014). Hydrological projections are, however, subject to considerable uncertainties, originating from the chain of models employed to generate the projections and giving rise to the cascade of uncertain (Wilby and Dessai, 2010). In this commentary, we discuss the link between climate projections and impact models and question whether the daily perspective prevailing in bias correction methods is appropriate for the modelling of future impacts. At the head of the impact modelling chains, general circulation models (GCMs) are run under different emission scenarios. These GCM simulations are performed on a comparatively coarse grid and have to be downscaled to a finer grid to improve the representation of atmospheric processes driving runoff generation at the catchment scale. For this purpose, regional climate models (RCMs) are often used and are run for a certain region on Earth, using the GCM simulations as boundary conditions. While RCMs provide simulations at scales similar to typical catchment areas, RCM simulations of temperature and precipitation often show considerable biases when simulations for current conditions are compared with observed data. Fischer et al. (2012), for instance, compared the simulation results of 20 GCM–RCM chains with gridded data based on observations (E-OBS) for the period 1980–2009 for three regions in Switzerland. They report, in particular, that the majority of the model chains provided excessively wet simulations for winter periods. The median of the 20 projections for mean winter precipitation was about twice as high as the corresponding gridded observations. Obviously, we cannot expect hydrological models to simulate reasonable runoff or other variables such as snowpack, when fed with so heavily biased climate simulations. Similar biases in GCM–RCM chains have been found in many studies in various regions and spatial scales, which led to the development of bias correction methods (Christensen et al., 2008; Piani et al., 2010; Dosio and Paruolo, 2011; Stoll et al., 2011). Various post-processing methods have been suggested to reduce the biases in simulated temperature and precipitation time series (reviews by Maraun et al., 2010; Teutschbein and Seibert, 2012). Although the influence of the bias correction on discharge simulation depends on the amplitude of the original bias, such bias correction generally leads to improved discharge simulations (Rojas et al., 2011; Muerth et al., 2013). However, the post-processed fields were shown to be critically sensitive to the assumptions about the future evolution of the biases (Buser et al., 2009) and to the observational data set used as a reference (Sunyer et al., 2013). Further, most bias-correcting methods do not account for the origins of the biases, often lying in the climate model formulation of highly non-linear processes (Maraun, 2012; Bellprat et al., 2013), which led to concerns about their ability to correct future biases in a robust way (Cloke et al., 2012; Ehret et al., 2012; Teutschbein and Seibert, 2013). In a sense, bias correction might provide a right answer (i.e., simulations looking like observations) but not necessarily for the right reasons. These limitations did not prevent the use of bias correction methods to grow in hydrological impact studies, essentially because of their ease of application and because without bias correction, GCM–RCM simulations can lead to substantial errors in hydrological projections. Among the different bias correction methods that have been suggested, quantile mapping has been found to provide particularly good results (Themeßl et al., 2011b; Teutschbein and Seibert, 2012) and is increasingly used (e.g., Finger et al., 2012; Forzieri et al., 2014). However, in almost all impact studies, the bias correction is performed for daily data. Discharge or other hydrological variables such as soil moisture or groundwater levels, on the other hand, often are a result of accumulated precipitation and/or temperature anomalies over several days, weeks or months. With this commentary, we raise the question whether the bias correction performed on daily data is appropriate for hydrological impact studies.

5.2 Beyond the daily perspective

Extremes in discharge, i.e., both high and low flows, are in most cases caused by precipitation events spanning more than a day and are strongly influenced by antecedent wetness conditions. This is clear for low flow conditions, which are not caused by a single day without rainfall or snowmelt but by a period of small water inputs. For instance, the heat wave that struck Europe during the summer of 2003 found its roots in the precipitation and soil moisture deficit accumulated since the spring of 2003 (e.g., Fischer et al., 2007). Similarly, high flows are usually caused by rainfall over several days rather than by a single day's rainfall. A recent example for Switzerland is the August 2005 flooding event, which was preceded by wet conditions and was driven by a multiday rainfall event, exceeding 260mm in 72 h in some parts of the Alpine range (Bezzola and Hegg, 2007; Jaun et al., 2008). For the purpose of this commentary, we explore in a simple way the sensitivity of discharge to multiday statistics of precipitation time series using rank correlation coefficients over 1986–2006. Our focus was on 11 mesoscale catchments (catchment area from 24 to 350km^2), all located north of the Alps in Switzerland with mean catchment elevations varying between 550 and 1400m a.s.l. The correlation between maximum annual discharge and the precipitation in the preceding n days (including the day of maximum discharge as the last day in the series) was computed. Additionally, the current precipitation index (C_{PI}, Smakhtin and Masse, 2000), which corresponds to the better known antecedent precipitation index (API) plus the precipitation on the current day, was computed iteratively as

$$C_{PI}(t) = kC_{PI}(t-1) + P(t) \tag{5.1}$$

where $P(t)$ is the precipitation on day t and the daily recession coefficient k was set to 0.85. The obtained C_{PI} was then correlated to the maximum annual discharge. The results indicated that for most study catchments, the maximum discharge was more strongly correlated to the precipitation summed over the preceding 2 to 6 days or to the C_{PI} than to the same day precipitation amounts (Figure 5.1). This supports the argument that multiday precipitation amounts are at least as important as 1-day amounts for the occurrence of floods. Here, we take the ability of the climate model to reproduce maximum 4-day precipitation amounts as an indication of its capacity to correctly capture multiday rainfall events. We assume that this variable should be represented correctly in the climatic data to allow for reliable flood modelling. We performed a similar analysis for low flows, correlating the lowest annual discharge with the precipitation amounts cumulated over the preceding weeks. Correlations were lower than for floods, principally because evaporation plays a larger role for low flows, but was not accounted for in the present analysis, which is focused on precipitation. The sensitivity of discharge to prolonged dry periods would be better studied using, for instance, a hydrological model instead of simple correlation coefficients. For this commentary, we assume that the minimum 14-day precipitation amount is relevant for the generation of low flows or situations with low soil moisture (hydrological or agricultural drought). Finally, we considered interannual precipitation variability, as these fluctuations are a driving component of climate time series (Blöschl and Montanari, 2010) and play a crucial role in the assessment of extreme events frequency and possible change thereof. To assess the interannual precipitation variability, we first computed the total precipitation amount in the winter and summer for each year of the reference period 1980–2009. For each season, we then computed the standard deviation among these 30 seasonal amounts and used it as an estimate of the interannual variability. The larger the interannual variability, the larger the chances of an anomalously dry or humid season to occur (for a discussion of the influence of climate variability on the occurrence of extreme events, refer to, e.g., Schär et al., 2004; Seneviratne et al., 2012b).

5.3 Biases before and after quantile mapping

We used the precipitation simulations by the RCM CLM forced by the GCM HadCM3Q0 run in the framework of the ENSEMBLES project (van der Linden and Mitchell, 2009). We compared these simulations with daily observations from a network of 392 rain gauges for the period 1980–2009. At least 80% of the daily observations were available for each station. The following six seasonal characteristics for the winter (DJF) and summer (JJA) were considered: mean precipitation, frequency of wet days (at least 1mm/day), mean precipitation on wet days, mean 14-day minimum, mean 4-day maximum and interannual variability. In contrast to the three first characteristics, the three last variables are multiday to interannual statistics. RCM interpolation was conducted using inverse distance weighting accounting for elevation dependence. Bias correction was performed using quan-

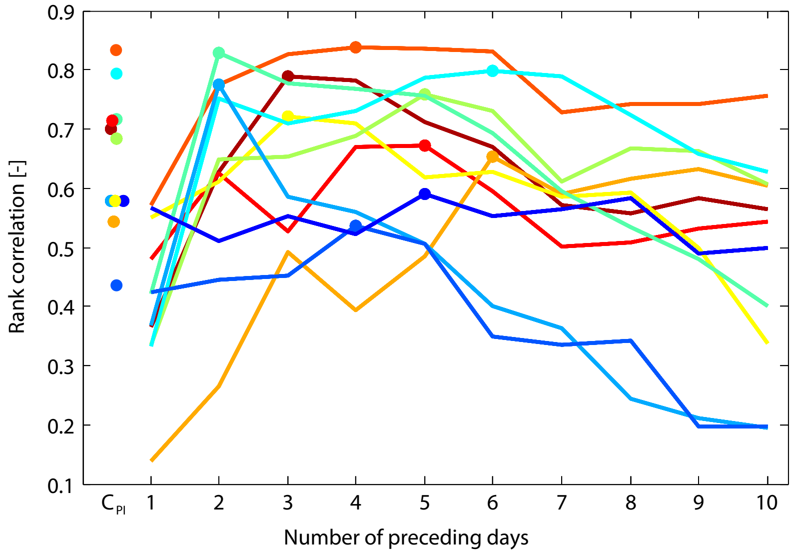

Figure 5.1: Rank correlation between the annual maximum discharge and the precipitation sum over n preceding days (including the day of maximum discharge) against the respective number of days. The strongest correlation is indicated by a dot. Additionally, the correlations between the annual maximum discharge and the current precipitation index (C_{PI}) are shown as dots at the left side of the figure. Each line and dot represents one of the catchments, which are all located north of the Alps in Switzerland. The colours indicate the respective mean catchment elevation, which varied between 550 (red) and 1400m a.s.l. (blue).

tile mapping relying on empirical cumulative distributions, which was found to be the most satisfying method in the comparison led by Gudmundsson et al. (2012). We evaluated the results by calibrating the bias correction on 27 years, evaluating it on the 3 remaining years and repeating the process a total of 10 times to cover the whole 30-year period. The obtained 30-year time series were then evaluated against station records. The difference in the variables computed over 30 years results from a combination of GCM–RCM biases, scale effects, measurement and interpolation errors and natural climate variability. Here, we do not separate these components and refer to their combined effect as bias in the following. The revealed large biases in all six characteristics (Figure 5.2, first and third columns). Winter precipitation is overall afflicted by large biases over the Alps, in particular over their northern flanks, with a large overestimation of the mean precipitation, as reported in other studies (Themeßl et al., 2011a; Fischer et al., 2012). Similar biases appear in the maximum 4-day total and in the interannual variability. The orientation and location of the zone with the highest bias suggest that the model does not correctly capture the influence of the Alps on precipitation. Summer precipitation is generally underestimated. After applying quantile mapping, biases for the first three characteristics were mostly eliminated (Figure 5.2, second and fourth columns). In contrast, for the three multiday to interannual characteristics, although the biases were generally reduced, large discrepancies between the observations and the RCM simulations remained after the application of quantile mapping. This result is in agreement with Wetterhall et al. (2012) who compared seven bias correction methods applied to 18 RCMs driven by ERA-40 reanalysis data in a UK catchment. They found that the correlation coefficients for the annual maximum 5-day precipitation were improved by the different bias corrections.

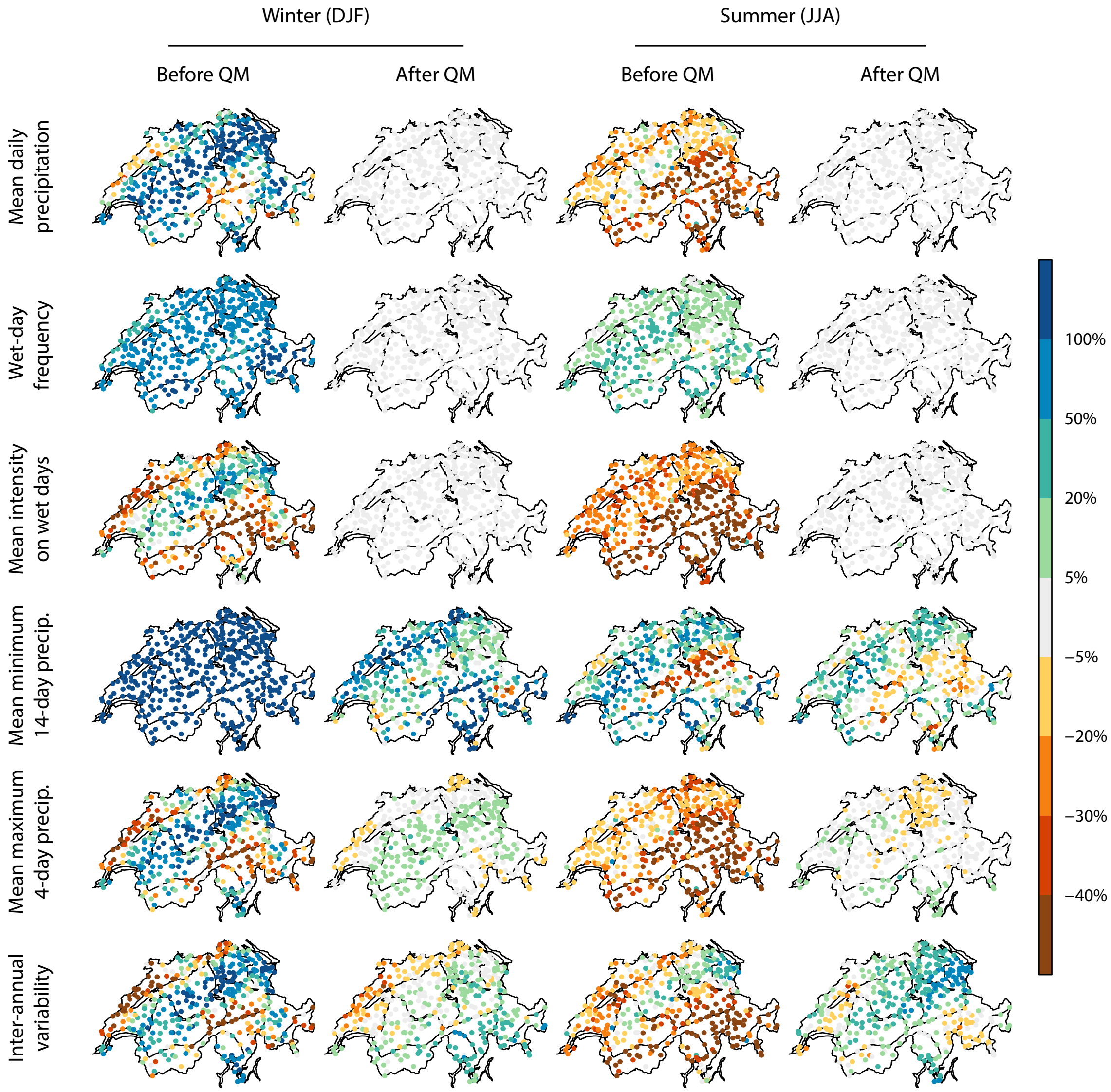

Figure 5.2: Comparison of RCM simulations for precipitation with observations from 392 gauge stations in Switzerland, each circle representing one station. Each map shows the relative bias between observations and RCM simulations for 1980–2009, the relative bias in variable X being computed as $(X_{RCM} - X_{OBS})/X_{OBS}$. The columns depict the effect of quantile mapping (QM) for winter and summer precipitation. The rows correspond to the selected characteristics of the precipitation time series. The main lakes and rivers, and hence valleys, are indicated in the background by black lines.

5.4 Robustness of quantile mapping

Just like the robustness of hydrological models that needs to be evaluated for events different from those used in the calibration period (Klemeš, 1986), we argue that bias correction methods should be evaluated on characteristics that they have not been explicitly been trained for. This enables us to establish whether bias correction approaches, such as the popular quantile mapping, are more than 'mathematical marionettes, dancing to match the calibration data' (Kirchner, 2006). The successful improvement of the first three characteristics in our study should not come as a surprise, as these characteristics can be directly derived from the cumulative distribution function (CDF) of the daily precipitation values. Quantile mapping does nothing more than train the CDF to agree with the observed CDF, and once the CDF is correctly reproduced, the three daily characteristics are corrected by definition. Fitting the CDF of each station is a simple task, given the number of degrees of freedom. The bias correction relies on 101 transfer functions (one per percentile) for each station and season. It thus would be surprising if the corrected daily variables did not correspond to the observed variables. In contrast to the daily characteristics, the multiday to interannual variables cannot be directly derived from the CDF of the daily precipitation amounts. Quantile mapping is hence not explicitly designed to correct for such statistics. Applying bias correction to daily values will change the multiday statistics, as daily and multiday statistics are to some extent related, for example, mean precipitation and 4-day precipitation amounts or wet-day frequency and minimum 14-day precipitation amounts. It is a priori unclear whether correcting the daily values would also lead to a correction of the multiday characteristics. Assessing this ability is an interesting robustness test for bias-correction methods. Our results revealed that this test was failed, as large biases in the multiday to interannual characteristics remained after the bias correction for most stations (Figure 5.2). Of particular concern is the remaining bias in the interannual variability, suggesting that the occurrence of dry and wet years is not represented properly in the data forcing impact models. In other words, bias correction via quantile mapping is a targeted and restricted correction. It helps to correct a set of characteristics, but the resulting simulations might still not agree with observed time series in other aspects. The importance of this depends on the sensitivity of the impact model. It is now unclear which biases in RCM projections are the most detrimental for hydrological projections, as systematic evaluations are so far lacking. The overestimation of the frequency of drizzle days by RCMs is, for instance, regularly reported in the literature (e.g., Teutschbein and Seibert, 2012) and can be efficiently corrected by quantile mapping (Figure 5.2). However, it might be that biases in other statistics, such as in the 4-day precipitation amounts or in the interannual variability, will induce more worrying errors from a hydrological perspective but are not satisfactorily corrected using quantile mapping.

While several studies investigate the effects of bias correction on precipitation and temperature (e.g., Piani et al., 2010; Dosio and Paruolo, 2011), those using impact models to investigate the consequence of bias reduction in the modelled impacts (e.g., Cloke et al., 2012; Muerth et al., 2013) are still rare. They are, however, of great importance, as simple sensitivity analyses such as the one used in this commentary do not handle the complex interactions between precipitation and temperature driving the annual cycle, droughts and floods. In contrast, using the post-processed time series to force a hydrological model would enable an integrated assessment of RCM and bias correction methods at the catchment scale.

5.5 Concluding remarks

To summarize, RCM simulations are usually corrected at daily time steps, and even after this correction, significant biases remain for hydrologically more relevant characteristics in the precipitation time series. While we here demonstrate this based on data for Switzerland, there is little reason to believe that the general result would be different for other regions. This is because this outcome is more related to the formulation of the bias correction scheme than to the climate or topography of the study area. Although some downscaling and bias correction method account for a wide range of timescales (Buser et al., 2009; Bordoy and Burlando, 2014), we argue that this issue has been largely overlooked in impact studies so far, and one has to be aware of this additional source of uncertainty.

Robust bias reduction could be achieved on the basis of a better understanding of the causes of biases in climate models. Recent studies, however, showed that large biases were found in the state-of-the-art climate model data set (Wang et al., 2014) and that some biases did not decrease and sometimes increased since the last model generation (Mueller and Seneviratne, 2014), despite considerable investments in model development and computing facilities. This suggests that identifying and addressing the causes of biases in climate projections will require considerable effort and time. In the meantime, and in parallel with further model development, we propose to further investigate bias- correction approaches by going beyond the daily timescale, as the number of impact studies relying on bias correction is increasing, both within and beyond the hydrology community. Finally, and from a more general perspective, we propose a more systematic quantification of the consequences of bias correction on impact simulations. Together with the traditional evaluation of biases in climate projections, this should provide new insights into the sensitivity of hydrological simulations to biases in the input data and guide the further development of impact modelling chains by, for instance, informing us on how much complexity should go into bias correction methods.

Acknowledgements

We acknowledge the EU-funded FP6 ENSEMBLES project and the Swiss Federal Office of Meteorology and Climatology for the climate simulations and the rain gauge observations. We thank Christoph Frei for the R package used to produce the maps. This work was funded by the Swiss National Science Foundation (grant 200021_131995).

Chapter 6

Trends in water balance as indicators of robustness for hydrological models in a changing climate

Nans Addor[1], Silviya Nikolova[1,2], Jan Seibert[1,3]

[1] Department of Geography, University of Zurich, Zurich, Switzerland
[2] Erasmus Mundus Master in Hydroinformatics and Water Management
[3] Department of Earth Sciences, Uppsala University, Uppsala, Sweden

In preparation

6.1 Introduction

Assessing the robustness of hydrological models under contrasted climatic conditions should be part of their evaluation (Klemeš, 1986; Andréassian et al., 2009), and there is a growing number of studies investigating this aspect of model performance (Seibert, 2003; Merz et al., 2011; Seiller et al., 2012; Coron et al., 2012, 2014). Robust models are particularly important when hydrological modeling is performed under future climatic conditions (Refsgaard et al., 2013), as models simulating well the system in its current state do not necessarily captures well its changes under climate change (see discussion on climate models in Racherla et al., 2012).

A pressing issue, directly relevant for model robustness under climate change, is the usually assumed stationarity of parameter values over time. Modeling experiments by Merz et al. (2011) and Coron et al. (2012) using conceptual hydrological models revealed that assuming transposability of parameters values over time in changing climatic conditions can lead to significant biases in discharge simulations. The amplitude of these biases usually increases with the difference between the climate (temperature and precipitation) of the calibration and simulation period. This raises the question whether parameter values need to be modified over time to reflect changes in hydrological processes induced by climate change. Such a question denotes a focus on the contribution of internal processes (i.e., catchment processes) to discharge generation.

Here we adopt a different perspective and explore the contribution of external forcings (i.e., changes in precipitation and temperature) to changes in discharge. After all, in a robust hydrological model, discharge variability should be induced by changes in the boundary conditions, and not by changes in parameter values. We propose to explore the robustness of hydrological models by assessing their ability to capture observed trends in the water balance (discharge and evapotranspiration). Trend analysis is commonly used for the detection of environmental changes induced by global warming, for instance changes in temperature (e.g., Ceppi et al., 2010). In hydrological research, trends in discharge observations for instance aim to investigate the response of watersheds to climate change (e.g., Stahl et al., 2010). For catchments in Switzerland, which is the region of interest of this study, Birsan et al. (2005) reported a significant increase of winter discharge in many catchments over the last decades. They suggest that it stems in particular from the increase of days with a minimum daily temperature above 0°C. Yet surprisingly few studies have so far tried to capture observed trends in discharge using hydrological models. Stahl et al. (2012) used for instance global hydrological models to assess trends in discharge in areas with a low density of catchments with long discharge records and little human impacts. They found that the mean of the ensemble captured satisfactorily the observed changes in discharge, but they did not explicitly evaluate model robustness when simulating the water balance under changing conditions.

Trends in water balance are a measure, integrated over time and the basin area, of the basin's response to climate change. Their analysis can highlight transitions from a hydrological regime to the other, such as the shift from nival to pluvial regime (or glacial to nival) caused by temperature increase and characterized in particular by an earlier onset of the discharge increase resulting from snow and ice melt (Barnett et al., 2005; Bundesamt für Umwelt BAFU (Eds.), 2012). Hydrological models driven by climate projections inform us about which changes to expect in the future, and as such can provide useful insights into changes already occurring, but not emerging yet from discharge natural variability. Further, although several studies looked at transient discharge changes at the annual scale (e.g., Merz et al., 2011; Coron et al., 2012), changes at shorter time scales can be more significant, as annual averages can for instance mask a decrease of summer and an increase of winter discharge.

Some of these seasonal changes can be captured by a hydrological model, even of low complexity, in which the value of model parameters is held constant. For instance, snow melt simulated by a degree-day routine will occur earlier under a warmer climate, without requiring any change in the routine. If the simulated trend corresponds to the observed one, the system may be described as robust or flexible under a changing climate. In contrast, a model might fail to reproduce observed hydrological changes when constant parameter values are assumed. In such situations, in which a model 'breaks', time-dependent parameter values might help to reflect the changes in hydrological processes.

To explore model robustness under changing conditions, and the potential benefit of time-varying model parameter values, we used the conceptual model HBV (Seibert and Vis, 2012) over 1971-2010 in four undisturbed research catchments in Switzerland. We calibrated HBV first over the whole period, and then for each of the four consecutive 10-year periods, similarly to Merz et al. (2011). We used these two setups, together with observations of daily precipitation, temperature and discharge to address the following questions:

1. What are the observed trends in temperature, precipitation, discharge and evapotranspiration in the study catchments over 1971-2010?

2. How well are trends in discharge and evapotranspiration captured by HBV using external forcing alone, i.e., with no change in parameter values?

3. When HBV is calibrated over 10-year periods, is there a drift in parameter values and if yes, can this drift be related to observed changes in physical processes or to calibration artefacts?

6.2 Data and methods

6.2.1 Study catchments

Our study investigates changes in the water balance induced by global warming, so we excluded catchments subject to human alterations, such as modifications of soil cover or the construction of dams. We restricted our attention to research catchments designated and monitored by the Swiss Federal Office for the Environment. Further the selected catchments had to satisfy the following criteria:

1. Daily discharge obeservations available since at least 1971.

2. No significant glacierized area, because glaciers add another source of uncertainty to the determination of the water balance, in particular because their contribution varies over time as they retreat.

3. Low percentage of karstic soils, as the particularly heterogenous storages and water flows they lead to is poorly captured by conceptual models calibrated using discharge data alone (e.g, Hartmann et al., 2013).

4. Catchments representative of the main hydrological regimes of Switzerland, except glacial regimes.

Key characteristics of the selected catchments are presented in Table 6.1. We retrieved temperature T and precipitation P data from the 2 km gridded datasets RhiresD and TabsD (Frei, 2013) produced by the Swiss Federal Office of Meteorology and Climatology, MeteoSwiss. Although HBV does not fully benefit from the 2 km resolution because of its semi-distributed nature, it is an advantage that stations data were combined with state-of-the-art methods, using nonlinear profiles and non-Euclidean distances for temperature, instead of e.g., simple IDW or bilinear interpolations.

Observations of daily discharge Q were provided by the Swiss Federal Office for the Environment. The periods used for the averaging of the time series are as follows. Winter is defined as the period from December to February and summer as June to August. Further, we consider hydrological years throughout this study, with for instance the year 1993 starting on October 1, 1992 and finishing on September 30, 1993. Similarly, winter 1993 started on December 1, 1992 and finished on February 28, 1993. All the annual

River	Gauging station	Area [km^2]	Mean elevation [masl]	Regime	Karstic area [%]
Dischmabach	Davos	43.3	2372	glacio-nival	0
Allenbach	Adelboden	28.8	1856	nival alpine	8
Murg	Waengi	78.9	650	pluvial-inferior alpine	0
Breggia	Chiasso	47.4	927	pluvio-nival meridional	0

Table 6.1: Key characteristics of the study catchments (www.hydrodaten.admin.ch).

averages, including for temperature and precipitation, were computed for hydrological years.

6.2.2 Hydrological model and modeling experiments

The hydrological model HBV is a conceptual semi-distributed model. The version used in this study (Seibert and Vis, 2012) was run using a discretisation of the catchments in 100 m-elevation zones. HBV has three main routines, namely a snow routine, a soil routine and a groundwater routine, and relies on 14 free parameters determined here by calibration. In order to asses eventual trends in the parameter values as a consequence of climate change, one key parameter for each routine was scrutinized:

1. snow routine: P_{CFMAX} [mm°C^{-1}d^{-1}] is the degree-day factor regulating snow melt and the refreezing of liquid water in the snow pack (Eqs. 1 and 2 in Seibert and Vis, 2012).

2. soil routine: P_{FC} [mm] is the field capacity (the capacity of the soil reservoir), which influences the partitioning of the available water (rainfall and snowmelt) between the soil and the upper groundwater reservoir, and is also involved in the estimation of the actual evapotranspiration (Eqs. 3 and 4 in Seibert and Vis, 2012).

3. groundwater routine: P_{K0} [d^{-1}] is the recession coefficient of an additional outlet of the upper groundwater reservoir (Eq. 5 in Seibert and Vis, 2012).

The potential evapotranspiration E_{POT} was estimated from extraterrestrial radiation and the mean daily temperature (Eq. 3 in Oudin et al., 2005). The actual evapotranspiration E (referred to as evapotranspiration in continuation) was estimated as the difference between precipitation and discharge ($P - Q$), which implies that changes in catchment water storage and the contribution of eventual glaciers to discharge are considered small enough to be neglected.

The mean precipitation and temperature prescribed for the whole catchment were adjusted for each elevation zone by assuming a precipitation increase of 5% per 100 m and a lapse rate of 0.5°C per 100 m, as the typically used value of 0.65°C per 100 m appeared to be too high in our area and in other mountainous ones (Blandford et al., 2008; Immerzeel et al., 2014).

To assess whether trends in the climate data are sufficient to reproduce eventual trends in observed discharge and estimated evapotranspiration, we performed two calibrations relying on different periods:

Calibration A: a calibration for the whole 40-year period, which is considered, for this setup and period, as the best calibration with constant parameter values.

Calibration B: a separate calibration for each of the four consecutive 10-year periods, which enabled the investigation of drifts in parameter values, and at a later point, will enable the assessment of the eventual benefit of time-dependent model parameters.

Calibration was performed using a genetic algorithm (Seibert, 2000). The objective function was an equally-weighted combination of Nash-Sutcliffe efficiency (Nash and Sutcliffe, 1970) and mean absolute relative error (as in e.g., Staudinger et al., 2011). The calibration algorithm was run 10 times for each period and catchment, yielding 10 parameter sets of comparable quality. Note that other similar studies (Merz et al., 2011;

Coron et al., 2012) considered a larger number of basins, but did not explore parameter uncertainty for individual basins.

6.2.3 Trends in the water balance and changes in HBV parameter values

Changes in temperature, precipitation, discharge and evapotranspiration were assessed using non-parametric Mann-Kendall tests, whose null hypothesis is that no monotonic trend (either upward or downward) is present, and are widely used in climate change studies (e.g., Schmocker-Fackel and Naef, 2010; Ceppi et al., 2010).

We tested the times series for significant trends, and assessed whether HBV simulations showed (or not) the trends we found (or not) in observations (estimations for evapotranspiration). In the following we consider that HBV performs well when the simulated time series shows a trend when the observations do, and no trend when there is none in the observations. The level of significance is 5%.

Discharge projections indicate that the future hydrological regime of a given catchment will be similar to the current regime of catchments at lower elevation (e.g., Bosshard et al., 2013b). It hence might be possible to apply a trading-space-for-time approach (Merz et al.,

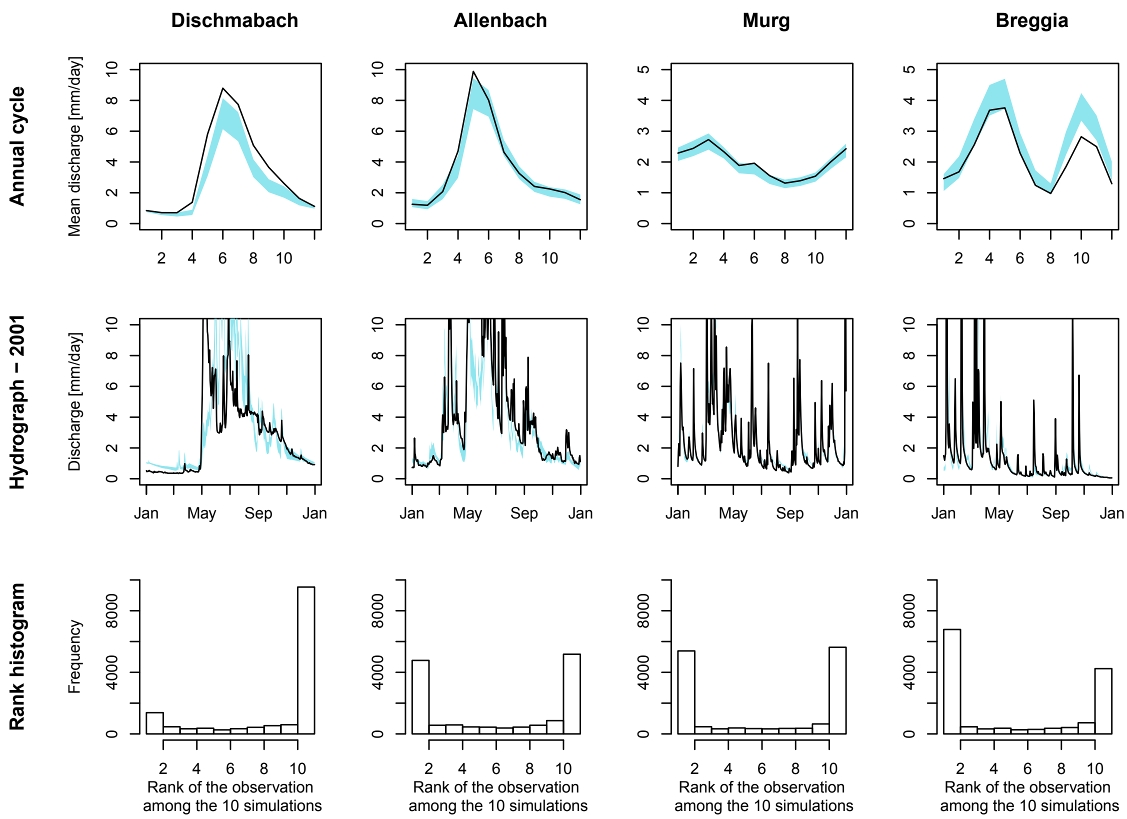

Figure 6.1: Observed discharge (black line) and spread of the 10 simulations (calibration A, blue area) for the 1970-2009 annual cycle (top row) and 2011 daily values (middle row). Rank histograms (bottom row) indicate the rank of the observed discharge among the 10 ensemble members for each day of 1970-2009. A U-shape means that observations are frequently outside of the ensemble range, a J-shape means that the ensemble usually underestimes the observations.

2011; Singh et al., 2011), and transfer parameter values from catchments at low elevation to catchments at higher elevation to model their future discharge. To test whether space can be traded for time in our study catchments, we produced a novel type of plot, and represented on the same x-axis different catchments and different time periods ordered according to their mean precipitation or temperature, and on the y-axis the value of a selected model parameter for these periods and catchments.

6.3 Results

6.3.1 Performance and uncertainty sampling of HBV simulations

Before investigating trends in HBV simulations, and discussing parameter values, it is necessary to investigate how well HBV captures the observed discharge, and to assess how much uncertainty is covered by the ensemble of simulations based on the 10 parameter sets generated by calibration A. The annual discharge cycle is well reproduced, the ensemble usually encompasses the observations, but underestimates the discharge in the Dischmabach catchment during most the year, as well as both peaks in the annual cycle of the Breggia (Figure 6.1). The hydrographs for 2011 and the rank histograms further reveal that for daily values, the ensemble is usually too narrow to include the observations. Increasing the number of parameter sets does not significantly change this situation (preliminary result, not shown). This underscores that parameter uncertainty is only one source of uncertainty, and that for instance errors in the input data, or the incorrect representation of catchment processes because of model formulation, are not reflected by the spread of the ensemble.

6.3.2 Observed and simulated trends in temperature, precipitation, discharge and evapotranspiration

Temperature increase is significant both for annual and summer means in the four catchments (Figure 6.2). It is weaker in winter (Ceppi et al., 2010). In contrast to temperature, seasonal and annual precipitation means do not show significant trends.

The correspondance between observed and simulated discharge trends is generally satisfying (Figure 6.3). HBV captures for instance the significant increase of winter discharge in Allenbach catchment, and does not generate significants trends in the other catchments, where the observations do not show significant trends either. The temperature increase in winter does not systematically lead to an increase of winter discharge, that one might expect as a result of more precipitation falling as rain, and rapidly contributing to discharge generation. This shift is clear in hydrological simulations driven by climate models, in particular for alpine (temperature-driven) catchments (Bosshard et al., 2013b), but so far does not systematically emerge from discharge records (Birsan et al., 2005). Although HBV relies on a conceptual representation of snow accumulation, based on temperature and not on energy balance, it captures this tendency well. Overall, the discharge evolution for Allenbach and Murg is in good agreement with the simulations. Although the annual Breggia discharge overestimated, the simulations capture well its observed decrease, because of a rather stable bias over time. Finally, as could be expected from the previous section, the spread of the ensemble of linear regressions is quite narrow and does not include the linear regressions derived from the observations.

None of the four catchments shows a significant increase in evapotranspiration, which is consistent lysimeter measurements in Switzerland (Figure 6l in Seneviratne et al., 2012a) and with the finding that evapotranspiration does not necessarily increase when temperature increases (Jung et al., 2010). Note that in the Dischmabach catchment, the observed discharge exceeds the observed precipitation. This might be partially explained by glacier melt, which we did not account for, but might also originate from an error in the observed time series derived for this catchment. These two possible causes contribute to explain why HBV underestimates Dischmabach discharge.

It is important to note that the quite good robustness of HBV stems partially from the setup of calibration A, which uses the whole dataset for parameter estimation. Simulations with parameters estimated using only one decade of the time series (calibration B) will probably be less robust.

6.3.3 Trading space for time

If parameter values could be directly inferred from the catchment mean temperature (precipitation), then there would be a clear organisation of the symbols in each panel of the first (second) row of Figure 6.4. For instance, if the value of a parameter increases when the temperature during its calibration period increases, then its value would be higher for lower (warmer) catchment. It is however hard to find such a clear pattern in the symbols of Figure 6.4. P_{CFMAX} decreases as the temperature in Allenbach increases, yet its value is lower in Dischmabach, although the Dischmabach is higher and has a lower mean temperature. Similarly, P_{FC} seems to increase with mean precipitation, but in Allenbach its value is much higher than what could be expected from its value in other three catchments. Finally, parameter equifinality is well represented by the spread of parameter values for the

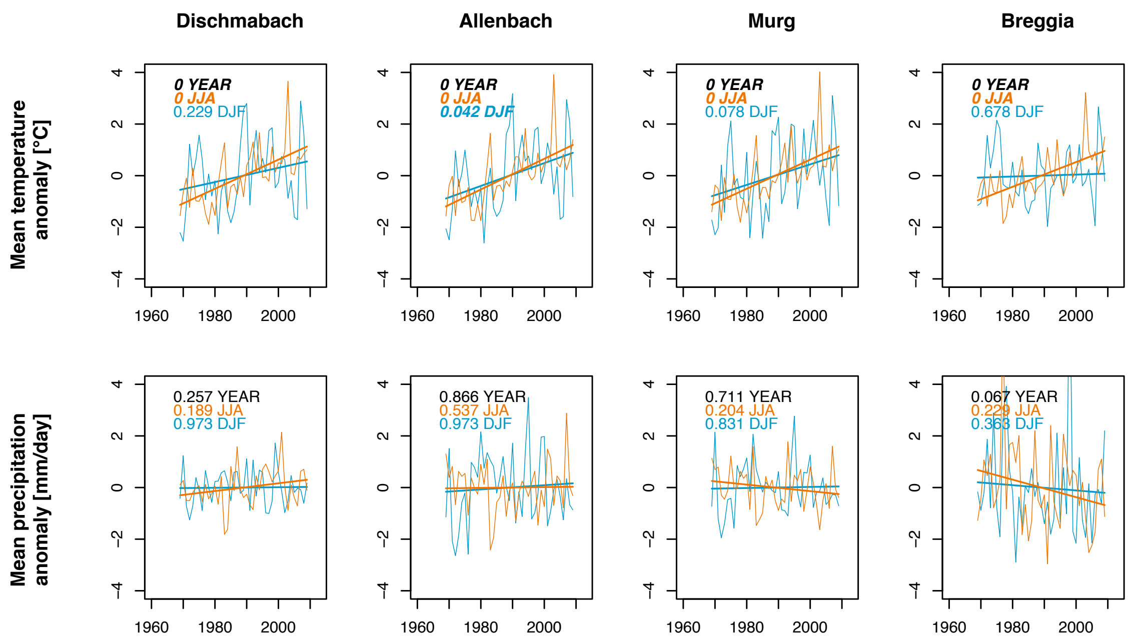

Figure 6.2: Variations of temperature (top) and precipitation (bottom) in the four study catchments in winter (blue) and summer (orange) for 1964-2009, and corresponding linear regression (thick lines). The p-value of the Mann-Kendall test for the seasonal and annual means is provided in the upper left corner of each panel, and is in bold if below 5%.

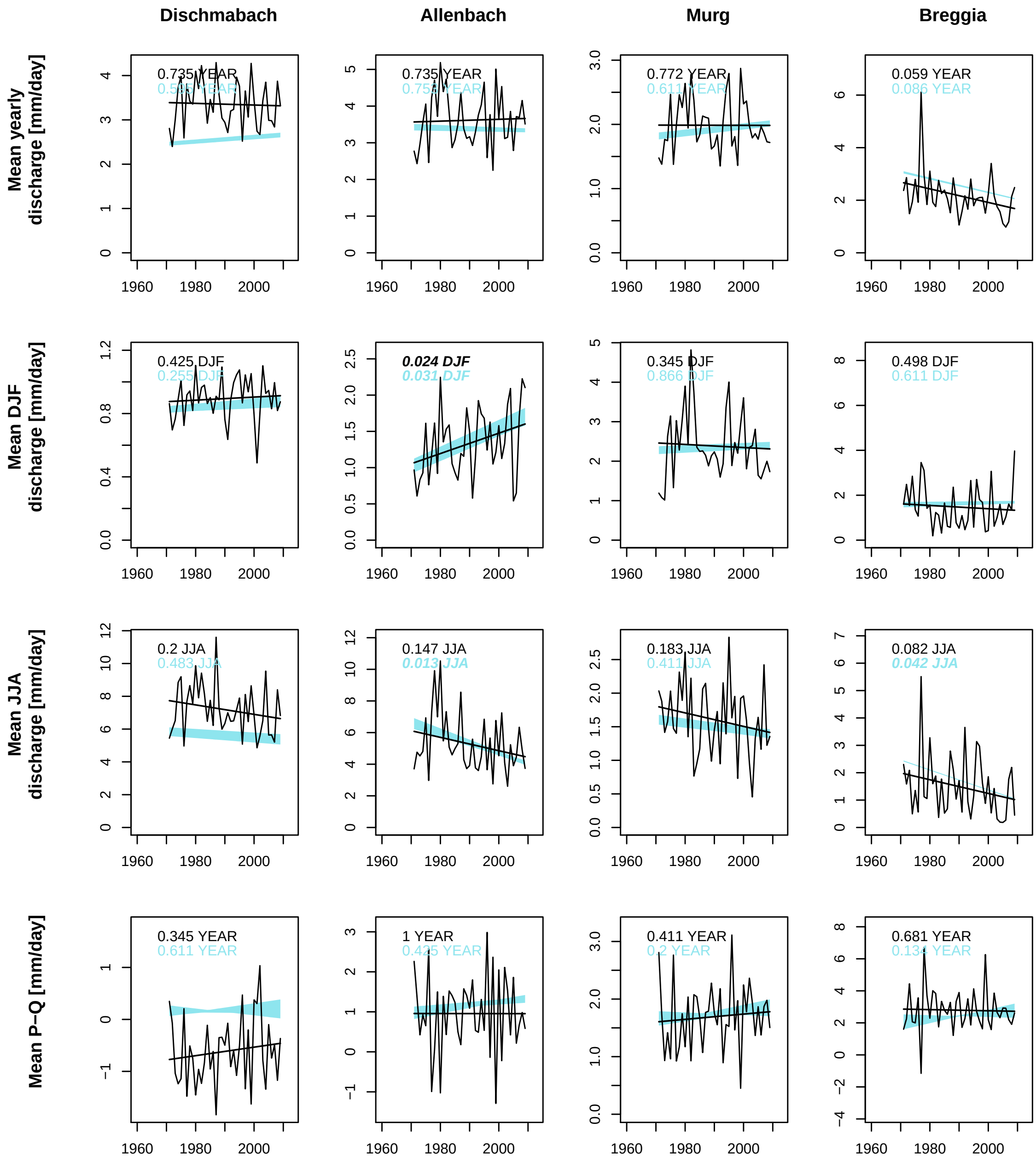

Figure 6.3: Variations of yearly (top row), winter (second row), summer discharge (third row) and yearly evapotranspirtation (bottom row) in the four study catchments for 1970-2009. Black lines denote observations and their associated linear trend. The blue areas are the envelopes of the trends simulated using the 10 parameter sets (calibration A). The p-value of the Mann-Kendall test for the observations and the mean of the ensemble is provided in the upper left corner of each panel, and is in bold if below 5%.

same period and catchment (same symbol and color), which underscores that for a given catchment under given climate conditions, parameter uncertainty is important.

6.4 Discussion

6.4.1 HBV robustness under changing climate conditions

The results obtained so far show that changes in the water balance are well captured by HBV even if its parameter values are kept constant. Upcoming simulations, using parameter sets from calibration B, will provide further insights into HBV robustness in changing conditions. Its robustness might decrease if parameter values obtained from the first (coldest) decade of the time series (1970s) are used for the simulation of the last (warmest) decade. Further, running HBV over the 40-year period using successively the four parameter sets derived from the 10-year periods will allow to test the benefit of a time-dependant parameters. It is possible that the gain will be modest. This would point towards structural problems, and advocates for the identification and further development of more robust model formulations.

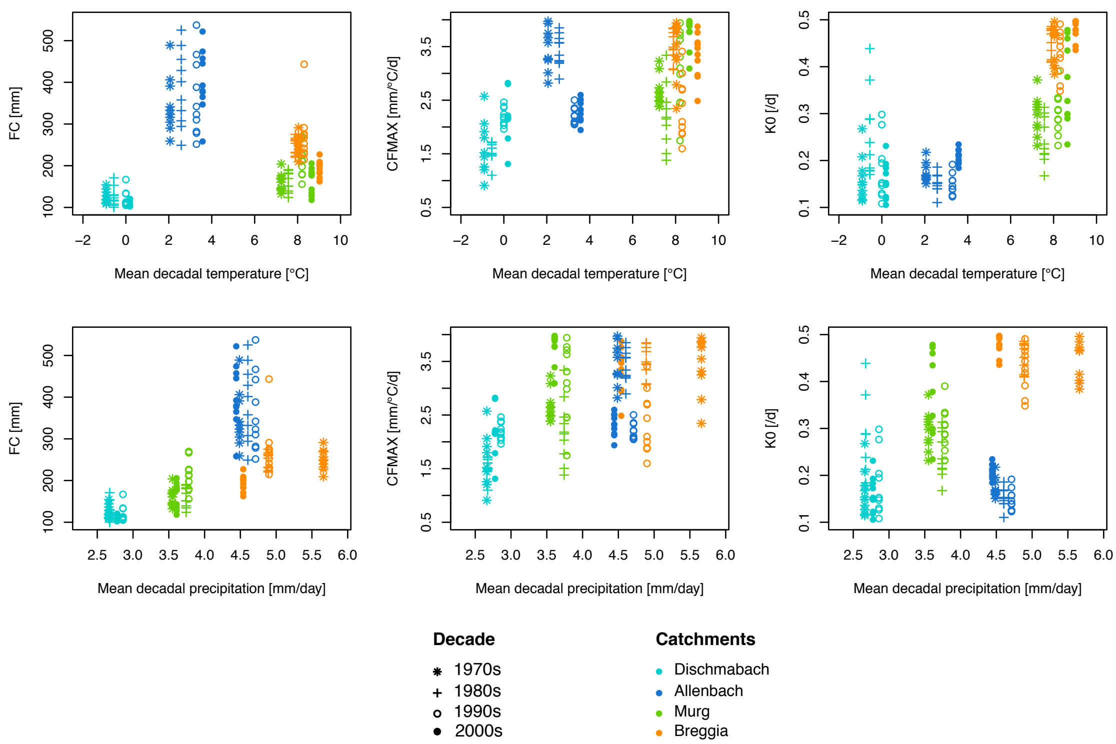

Figure 6.4: Parameter values for P_{FC}, P_{CFMAX} and P_{K0} (first, second and third column) estimated using calibration B. Each element corresponds to a parameter set, the color refers to the catchment and the symbol to the period of the calibration. 10 parameter sets were produced for each catchment and period. The x-coordinate of each symbol indicates the mean temperature (top) or precipitation (bottom) during its calibration period.

6.4.2 Relating time and space to increase the robustness of parameter estimation?

Encouraging results were already obtained by trading time for space to estimate parameter values (Singh et al., 2011). Yet as shown in the framework of the decade of Predictions in Ungauged Basins (PUB, Hrachowitz et al., 2013), the transfer of parameters from one catchment to the other is not trivial. In fact, before considering the change of parameter values over time, we need to ask ourselves: what do parameter values correspond to? They reflect hydrological processes, which depend in particular on the climate, vegetation and geology of the catchment, hence on catchment properties. But they are also 'effective parameters' (Beven, 1989): they help to capture a wide range of physical processes leading to discharge, but they are not themselves parameters with a physical meaning, which makes them difficult to transfer from one catchment to the other. Further, parameter values help to compensate for errors in the input data, but since these errors are often catchment specific, different catchments need different corrections. Finally, parameter values often also help to correct for the errors introduced by the inadequate formulation of hydrological processes in models. Merz et al. (2011) for instance argued that evapotranspiration would increase because of the significant increase in air temperature, and used this argument to explain drifts in parameters values. But instead it could be that the parameterisation chosen for evapotranspiration is too sensitive to temperature increase (Sheffield et al., 2012), overestimates real evapotranspiration under warm conditions, and creates a compensating drift in parameter values, which would be a modeling artefact and not a sign of catchment response to climate change.

6.5 Preliminary conclusions and outlook

We showed that the evolution of the water balance observed over the last four decades in four Swiss catchments can be captured reasonably well by HBV, even if its parameter values are kept constant. This is a sign of robustness not necessarily expected from a simple conceptual model. Upcoming simulations, using shorter periods for model calibration, are however expected to be less robust than the simulations showed in this study. They will enable us to assess the potential benefits of time-dependent parameter values for hydrological modelling in changing climatic conditions.

Chapter 7

From products to processes: Academic events to foster interdisciplinary and iterative dialogue in a changing climate

Nans Addor[1], Tracy Ewen[1], Leigh Johnson[1], Arzu Çöltekin[1], Curdin Derungs[1], and Veruska Muccione[1]

[1] Department of Geography, University of Zurich, Zurich, Switzerland

Published in Earth's Future, 2015
doi:10.1002/2015EF000303

Abstract

In the context of climate change, both climate researchers and decision makers deal with uncertainties, but these uncertainties differ in fundamental ways. They stem from different sources, cover different temporal and spatial scales, might or might not be reducible or quantifiable, and are generally difficult to characterize and communicate. Hence, a mutual understanding between current and future climate researchers and decision makers must evolve for adaptation strategies and planning to progress. Iterative two-way dialogue can help to improve the decision making process by bridging current top-down and bottom-up approaches. One way to cultivate such interactions is by providing venues for these actors to interact and exchange on the uncertainties they face. We use a workshop-seminar series involving academic researchers, students, and decision makers as an opportunity to put this idea into practice and evaluate it. Seminars, case studies, and a round table allowed participants to reflect upon and experiment with uncertainties. An opinion survey conducted before and after the workshop-seminar series allowed us to qualitatively evaluate its influence on the participants. We find that the event stimulated new perspectives on research products and communication processes, and we suggest that similar events may ultimately contribute to the midterm goal of improving support for decision making in a changing climate. Therefore, we recommend integrating bridging events into university curriculum to foster interdisciplinary and iterative dialogue among researchers, decision makers, and students.

7.1 Introduction

'There are many approaches to conceptualising uncertainty' – *Pete Fisher*

Uncertainty is an inherent component of research on climate change and of adaptation to its impacts. It spans from projecting future climate change (e.g., Hawkins and Sutton, 2009) to assessing regional impacts and vulnerabilities (e.g., Füssel and Klein, 2006; Bosshard et al., 2013a) and designing adaptation policies (e.g., Dessai and Hulme, 2004). Despite the presence of uncertainty throughout these fields, both its sources and the methods for its handling are heterogeneous. Uncertainty regarding climate change is often conceptualized as a cascade rather than a singular problem (Wilby and Dessai, 2010), with uncertainties from the climate system at the top of the pyramid cascading to lower levels representing impact modeling and then adaptation (the so-called top-down approach). Moving down this chain, uncertainties that principally stem from physical processes and can be quantified (e.g., Knutti and Sedláček, 2012) may compound with uncertainties of a social, political, or economic nature, which are often more qualitative and less quantifiable (e.g., Demeritt et al., 2007). Incorporating these uncertainties into decision processes is challenging, in particular, because of the different nature of these uncertainties; the distinct structures in which climate researchers, impact modelers, and decision makers operate (e.g., Dabelko, 2005; Vogel et al., 2007); and the different conceptions of these actors regarding what information is useful for decision making (Dilling and Lemos, 2011). In an attempt to overcome these limitations, scholars have advocated a better consideration of end-users' vulnerabilities before incorporating climate information (the so-called bottom-up approach, e.g., Brown and Wilby, 2012). Finally, in order to progress with decision making in the presence of uncertainties, there is a growing body of literature suggesting that scientists, policymakers, and concerned publics should go beyond one-way approaches (either top-down or bottom-up) and instead engage in interdisciplinary and iterative dialogue, hereafter referred to as IID (e.g., McNie, 2007; Dilling and Lemos, 2011).

To foster IID, several channels have been proposed and are now operational. They include boundary organizations, climate services agencies, and informal knowledge networks, which we describe below. There is, however, little discussion in the literature about how people involved in these institutions gain the understanding and skills necessary to help IID occur. In this commentary, we propose to develop academic events with the goal of familiarizing researchers, decision makers, and students to IID. We use outcomes from a workshop we organized to reflect on the following question: how can academic events set the basis for interdisciplinary and iterative dialogue to occur?

7.2 Top-down and bottom-up approaches for decision-making under uncertainty

How are we to deal with uncertainties when producing climate projections, assessing impacts and vulnerabilities, and designing adaptation strategies? Two main approaches can be distinguished, usually referred to as top-down and bottom-up approaches (see Weaver et al., 2013, for a comparison). In the prevailing top-down paradigm, the backbone consists of a model chain, usually involving one or several emissions scenarios and climate models, often followed by one or several downscaling methods. Downscaling derives locally relevant climate data from the global-scale predictions generated by coarser resolution climate models, which are used to drive one or several impact models at a finer scale. At each step, uncertainties are sampled using different models and/or parameter values and are then propagated to the next element of the model chain. Results are then presented as an ensemble of equally weighted model runs or combined as probability distributions (e.g., Tebaldi et al., 2005; Knutti et al., 2010). Climate change impacts are then often derived from the combined effect of several parameters (e.g., Fischer and Knutti, 2012) and are typically assessed at the regional scale (e.g., Addor et al., 2014). To advance our understanding of current and future changes, there is a steady effort in the research community to increase the complexity of climate and impact models. Yet although newer generations of models are better at representing the observed climate (e.g., Knutti et al., 2013), this does not necessarily lead to decreased uncertainty in the climate projections (Knutti and Sedláček, 2012). Further, part of the uncertainty is irreducible due to the natural variability of the climate system (Deser et al., 2012a) and would remain even in the hypothetical case of unlimited computing resources and deterministic knowledge of the system. Finally, if such top-down projections inform us about *which* changes to expect, researchers have increasingly questioned whether they provide much guidance on *how* to mitigate these changes (Dessai and Hulme, 2004; Dessai et al., 2009; Prudhomme et al., 2010; Brown and Wilby, 2012). As phrased by Lemos and Rood (2010), *useful* projections, i.e., projections that scientists perceive to be relevant for user groups, are not necessarily *usable* by these user groups, i.e., do not necessarily help them advance a decision process.

In response to this mismatch, some physical and social scientists have developed so-called bottom-up approaches. They build on the premise that system sensitivities and user needs and vulnerabilities must be understood first, and that climate projections should be used later to inform rather than to drive the analysis. Indeed, it is not critical that uncertainties be reduced or fully characterized, but rather that their effects on decisions taken are better understood in order to inform the decision making process (Brown and Wilby, 2012; Weaver et al., 2013).

Such approaches commonly rely on three main steps. One starts with the evaluation of the key sensitivities of a systemor the vulnerabilities of a particular population or a community (Wisner et al., 2004; Brown et al., 2012). Direct dialogue with stakeholder and user groups is essential to establish which circumstances might alter their particular activities. For instance, in the study by Brown et al. (2011), stakeholder groups whose activity is linked to Lake Superior (e.g., involved in commercial shipping or wastewater management) were asked under which lake levels and for what duration they would consider their situation as either 'acceptable', posing 'significant negative impacts, but survivable', 'intolerable without policy changes'. Other climate-sensitive systems include health and food supply, ecosystems, and infrastructure, with differential impacts depending upon a group's access to resources. As the next step in the assessment, researchers determine the climatic conditions (typically changes in temperature and precipitation) that would lead to critical situations, for instance using a stochastic framework (Steinschneider and Brown, 2013). Finally, climate information is used to determine how likely these situations might become in the future. Hence, climate change information only enters in the last stage of the approach, and although this may include climate model projections, it could just as easily include information derived from paleo-climate data or expert elicitations (Brown et al., 2012).

Bottom-up approaches offer several advantages. Since they involve users in the first stages of the study (i.e., including in problem definition, choice of scope, and selection

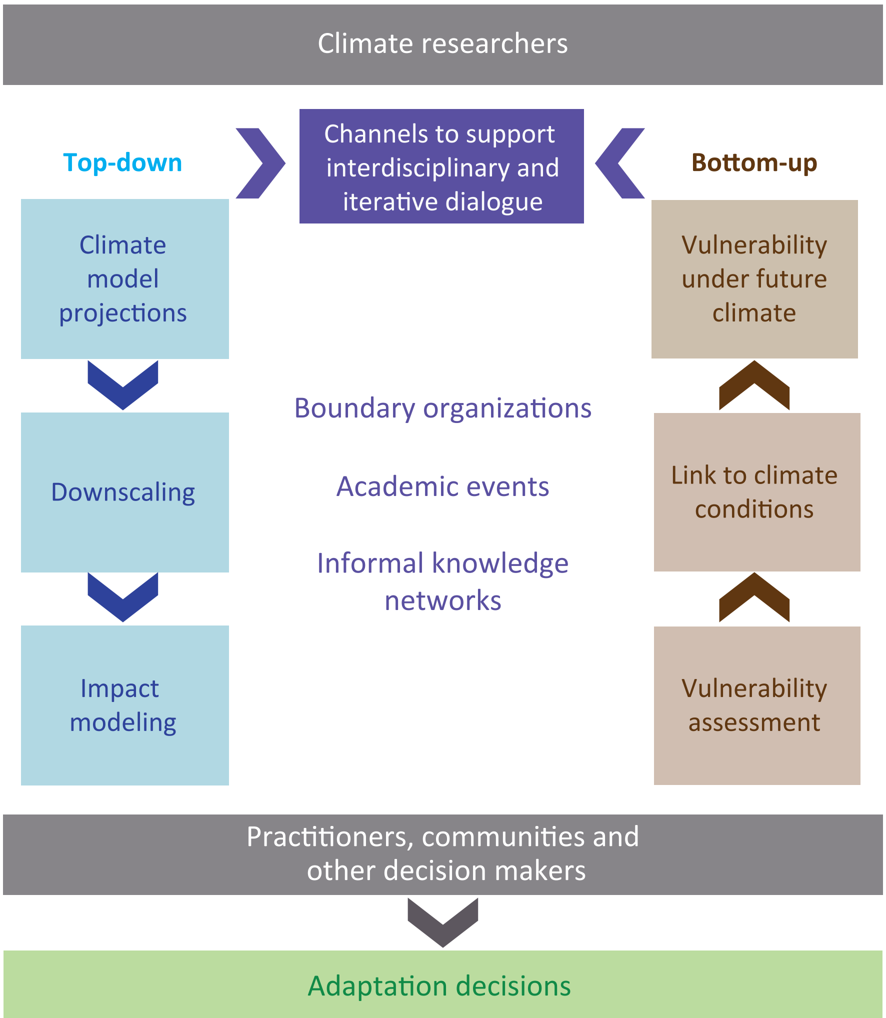

Figure 7.1: Schematic representation of the flow of information between climate researchers and decision makers in the top-down and bottom-up approaches (left and right columns, respectively). As these two approaches may turn into one-way communication modes, alternative approaches relying on channels to support iterative dialogue (as shown in purple, see text for examples) have been proposed to aid knowledge production and support decision-making under uncertainty.

of credible climate information), they are expected to increase the legitimacy accorded to climate projections, improve the relevance of outputs, and ultimately raise the chances of success of adaptive action (Vogel et al., 2007). They can contribute to the emergence of more robust decisions, i.e., solutions that perform well under a wide range of climate outcomes (Dessai et al., 2009; Wilby and Dessai, 2010). Further, bottom-up approaches respond to calls from social scientists for climate change research to engage with contextual or 'starting-point' vulnerability rather than with outcome vulnerability (Kelly and Adger, 2000). In the latter, vulnerability is an end-result of projected impacts on a particular exposure unit; in the former, it is the present incapability to cope with a variety of external changes as a result of interacting social, political, economic, and environmental conditions (Wisner et al., 2004; O'Brien et al., 2007). The bottom-up approach is better suited to integration with contextual vulnerability assessments such as community-based self-assessment of coping capacities and participatory risk-mapping exercises (Smith et al., 2000; Tschakert, 2007). Indeed, this approach significantly expands the definition of who is considered to be a relevant 'decision maker' to include community-level organizations that are too often left on the receiving end of decision making. If a vulnerability analysis reveals that the system or community is not particularly sensitive to changes in climate, or that the impact of other stressors far outweighs climatic impacts, then model projections may be irrelevant to the decision making and their time-consuming processing may be avoided (Brown and Wilby, 2012). Similarly, once vulnerabilities to changes in climate are established, new climate projections may be compared to these vulnerabilities as they become available, without the need to rerun the impact model (Prudhomme et al., 2010).

As discussed earlier, top-down approaches typically fall short when the information provided by climate researchers does not correspond to the needs of decision makers. Yet adopting a bottom-up approach does not automatically lead to an improved understanding between the parties. In fact, similar communication problems can occur, in which the expectations of the user groups are beyond the reach of the climate community. For instance, in a bottom-up framework, a sensitivity analysis might be carried out to determine under which climate conditions the reliability of a water supply system will be compromised. Although these conditions might be clearly identified, quantifying their probability under a changing climate can be at the very edge of, and often beyond, the present understanding of the climate system under global warming. Uncertainty is higher at smaller spatial scales and is higher in variables such as the variance than it is in the mean. Yet it is often precisely the changes in complex variables at small spatial scales that are relevant for impact studies and decision making (Brown and Wilby, 2012). One-way bottom-up requests from practitioners to climate scientists can end up in a dead end, if it turns out that available climate data and our present understanding of the climate system cannot provide the required information. In such cases, bottom-up approaches fall short. In summary, a key issue is the typical lack of convergence between the information that can be derived from climate data and the information needed to support decision making.

7.3 Fostering IID to aid decision making under uncertainty

To overcome the limitations of purely one-way approaches, scholars have advocated an increased two-way dialogue. Brown and Wilby (2012) propose to adopt a 'top-down meets bottom-up' framework, McNie (2007) calls for a 'reconciliation of the supply of scientific information with users' demands', and Dilling and Lemos (2011) argue that 'co-production of knowledge requires iterativity between scientists and potential users/stakeholders'. Similarly, Nowotny et al. (2001) argue that in order to generate knowledge that remains valid outside the confines of purely theoretical and experimental science, new modes of communication subject to 'frequent testing, feedback, and improvement' are required. Two key characteristics are crucial to sustain future dialogue: interdisciplinarity and iterativity. Interdisciplinarity should be understood here in a broad sense, such that it is not restricted to exchanges across academic disciplines, but instead includes exchanges between climate researchers and decision makers. Note that we conceptualize decision makers in the broadest terms, including practitioners within organizations, end users of projections, communities considering adaptation options, as well as more traditional policymakers and planners. Furthermore, dialogue between these actors should also be iterative, that is, not only two-way but also ongoing and open-ended. For such exchanges to maintain salience, credibility, and legitimacy for multiple audiences and actors, 'true dialogue' requires that 'scientists and users be brought together with equal standing for setting agendas, designing products, and evaluating successes' (Cash et al., 2006). We follow Lemos and Morehouse (2005) definition of iterativity as '(a) the extent to which the interactions between scientists and stakeholder participants influence how scientists pursue science and how stakeholders understand the possibilities and limits of science, (b) the range of uses to which the scientific knowledge may be put, and (c) the practical value of such knowledge'. Such iterative modes of knowledge production about climate change contain greater possibilities for innovation and societal impact (Lemos and Morehouse, 2005).

To further support improved dialogue and better accommodate uncertainty, various channels exist that help to foster interdisciplinary, iterative and more complex (multi-party) forms of dialogue (Bidwell et al., 2013; Hoppe et al., 2013). These include channels such as boundary organizations, academic events, and informal knowledge networks (Dilling and Lemos, 2011; Lemos et al., 2014, see Figure 7.1). Boundary organizations, for instance, help to foster dialogue by providing a network to broker information between scientists and practitioners with a focus on the science–policy interface (Lemos and Morehouse, 2005; Vogel et al., 2007). Additional channels to support dialogue and the flow of information include publicly funded projects or branches within federal organizations that have been launched to guide adaptation strategies, as well as university centers that help to bridge the gap between researchers and practitioners (e.g., UKCIP at the University of Oxford, Stockholm Environment Institute, Oregon Climate Change Research Institute, Pacific Climate Impacts Consortium at the University of Victoria, and African Climate and Development Initiative at the University of Cape Town, among many). Other informal information sharing takes place, for instance, on internet platforms that disseminate information on adaptation, allowing users to both access and share information and data (e.g., http://climate-adapt.eea.europa.eu, http://weadapt.org/), and via courses organized to guide the use of climate model output and appropriate use of downscaled projections for adaptation and policy development (e.g., using regional climate model data for Alpine impact research (Salzmann et al., 2013) and CSAG Winter School, University of Cape Town). This type of information sharing during a course or workshop can often have more impact and result in better application of the information than if learned elsewhere (e.g., via internet or journal article, Bidwell et al., 2013).

Despite the growing number of channels facilitating communication and dialogue between climate researchers, users, and practitioners, critical gaps still exist. In particular, many forums for exchange primarily engage already-established researchers and decision makers within traditional networks. Communities of climate researchers and decision makers are also evolving rapidly and new models that facilitate communication will need to take this into consideration, for example, through the development of bridging organizations that act to support and strengthen independent smaller networks and decentralize the flow of information, as is being done with the Great Lakes Integrated Sciences and Assessments Center (Bidwell et al., 2013). We propose that one way to expand both the range of participants and the content included in this dialogue, and thereby improve traditional communication channels, is through training and pedagogical responses that introduce students and academics from a wide range of backgrounds to interdisciplinary dialogue (shown in Figure 7.3).

7.4 A workshop on 'Uncertainty in Decision-Making in a Changing Climate'

We introduced an innovative workshop-seminar series at the University of Zurich to demonstrate the form such a response might take. Our workshop is used here as an example of how such events can be used to facilitate current and future dialogue between evolving groups of climate researchers and decision makers. The event was designed to generate interactions between a wide range of participants: nine expert speakers from academia, industry, government, and humanitarian aid and development, as well as bachelor, master, and PhD students (i.e., undergraduate and graduate students) and academic staff. The main goals were to:

1. provide participants with an overview of the current research on uncertainty and on how uncertainty is dealt with by decision makers,
2. overcome existing barriers to communication (e.g., limited opportunities for informal face-to-face interactions) and thereby enhance mutual trust and understanding on which collaborations can be based (Dabelko, 2005; Vogel et al., 2007), and
3. expose students at an early stage of their professional life to multidisciplinary collaborations (e.g., Gornish et al., 2013) and real-world problems involving decisions under uncertainty.

The event began with a 2-day workshop and was followed by case study assignments in which participants spent 2 months grappling with uncertainties with implications for decision making. This led to the incorporation of the material into an existing course that is mandatory for all masters students in the Department of Geography. More details on the workshop-seminar series are provided in the Supporting Information.

7.5 Shifting conceptions about decision-making under uncertainty

To explore the impact of our 2-day workshop on participants' perspectives about communication between scientists and decision makers, we conducted an anonymous opinion survey before and after the workshop. Participants were asked *'What information and tools should be exchanged between researchers and decision makers to better address uncertainty?'* The responses visualized in Figure 7.2 show a shift in perspectives regarding the relationship between researchers and user groups as a key outcome (see Supporting Section 7.7.2 for methodology). This shift is from a pre-workshop vertical conceptual model of interactions between researchers and user groups to a post-workshop horizontal model. Before the event (Figure 7.2a), participants' responses paralleled the dominant approach of top-down climate change impact modeling, prioritizing outputs that could be generated by academic researchers based on their expertise and their apparently more direct access to climate information. For instance, better data visualizations, quantifications of uncertainty, and metadata were among the most recurrent key words and were each suggested in approximately 20% of responses. In this vertical model, researchers bestowed data-based products to decision makers, who were minimally involved, by specifying the kinds of information they require. This model resembles the classic 'pipeline model' or 'loading-dock approach' of science–society relationships identified and critiqued by social studies of science and expertise (Nowotny et al., 2001; Cash et al., 2006), and indeed some respondents even used the term 'pipeline'.

In contrast, participants' responses after the event (Figure 7.2b) tended to prioritize *processes* rather than *products*. For instance, previously prioritized outputs such as visualizations were virtually discounted (appearing in only one response), and recommendations to produce quantifications of uncertainty and metadata were only half as frequent as before the event. Instead, the most common recommendations were for dialogue (35%) and more frequent and improved channels for communication (30%). In this 'flatter' model, researchers and decision makers engaged in institutionalized dialogues and frequent communication, exchanging their needs, expertise, and even personnel. There was an increasing recognition that such exchanges must allow for decision makers to specify and iteratively define what kind of information they need (mentioned in 20% of post-workshop responses

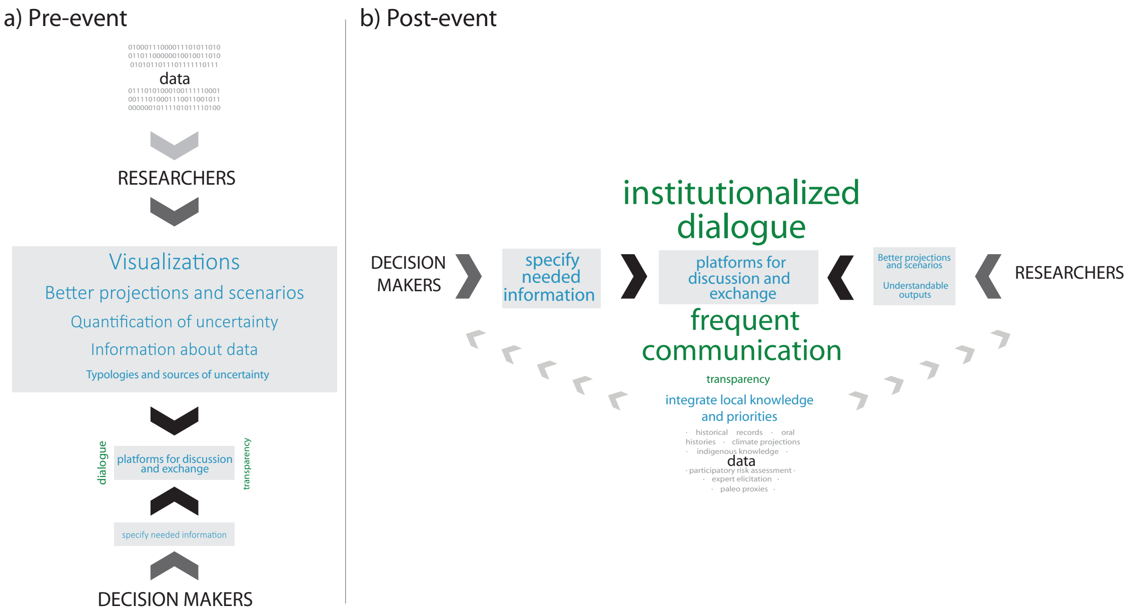

Figure 7.2: Shifting perspectives on the communication of uncertainties in the decision-making process, before and after an organized two-day workshop promoting dialogue exchange. Here, results from an opinion survey conducted before and after the workshop (a: n=53, b: n=42) are visualized, where font size is scaled based on the code's relative frequency in each set of responses. Items in blue are concrete outputs or deliverables, while items in green are processes; note the shift from the former to the later, from pre- to post-event responses.

versus 10% of pre-workshop responses). Responses also suggested that the local knowledge and priorities of affected communities struggling with the uncertainties surrounding climate change should inform both researchers and decision makers and shape the design and coproduction of relevant outputs – a concept that was completely absent from pre-workshop responses but appeared in 15% of post-workshop responses. This conceptual model of interaction, which we characterize as 'horizontal', is neither a strict bottom-up nor a top-down relationship. Rather, it is an iterative and flexible relationship that may assume different forms based on the needs, priorities, and data available.

A second set of qualitative evidence corroborates this perceptual shift among participants. The workshop ended with a spirited round table discussion and budgeting exercise that included the speakers and the audience. All participants were asked *'What are the most urgent cross-cutting challenges posed by uncertainty in a changing climate, and how can researchers and user groups collaborate to address them?'* To gage responses, we deployed a mock 'participatory budgeting' exercise in which participants were asked to allocate hypothetical grant money among eight proposals, derived from speaker presentations and discussions throughout the workshop (see Supporting Section 7.7.3 and Supporting Figure 7.4). Participants' budget allocations prioritized proposals that facilitated processes (e.g., dialogue, developing novel communication channels, acknowledging and engaging local knowledge) over proposals that resulted in end-products (e.g., better projections, better data, or insurance). These responses implied a shift away from a 'deficit model' of science communication and decision making, in which members of the public are conceptualized as simply lacking knowledge that scientists should produce and provide (Locke, 1999; Crow and Boykoff, 2014), and toward an iterative model of problem definition and research design that engages with multiple forms of knowledge, including local and indigenous 'uncertified' expertise about environmental change (Wynne, 1996; Collins and Evans, 2002; Vogel et al., 2007). They also reflected a recognition that the irreducible nature of some types of uncertainty means that social values and priorities are inevitably present in environmental decision making.

It is important to note that neither the two-round opinion survey nor the participatory budgeting activity were controlled experiments, but rather workshop exercises. As such, our results do not permit us to definitively assign causality for the patterns observed. As demonstration research, they do allow us to hypothesize that such academic events can lead to perceptual shifts about how best to address uncertainty in a changing climate, away from prioritizing 'pipeline' top-down models toward more horizontal, iterative, and interdisciplinary dialogue.

7.6 Conclusions and outlook

Uncertainty will continue to be an inherent part of climate change research and pose challenges for decision making. When dealing with uncertainty in the context of impact modeling, vulnerability assessment, or adaptation planning, it is crucial to better understand the effects of uncertainties on the decisions in question. IID is therefore central, as it enables bridging of current top-down and bottom-up approaches, thus overcoming existing barriers to communication and enhancing mutual trust and understanding on which collaborations can be based (Dabelko, 2005; Vogel et al., 2007). In this study, we explored how an academic event may set the basis for IID to occur.

We conducted two short evaluations: a two-stage survey and mock budgeting exercise. Their results suggest that even a short 2-day workshop can change participants' perspectives on addressing uncertainty. In both exercises, participants prioritized processes over products. Given these findings, we suggest that the workshop helped the participants to better conceptualize the myriad constraints of data, actors, and institutions attempting to address uncertainty. We formulate three hypotheses that could be systematically tested in future studies of similar events:

1. The workshop helped the participants to better understand the sources of uncertainty inherent to climate projections, to acknowledge that these uncertainties are not necessarily reducible (natural variability), and hence to realize the importance of working with decision making schemes that can accommodate these uncertainties.

2. The workshop helped the participants to recognize that beyond the quest for better models and less uncertain projections, it is crucial to achieve a better correspondence between the information provided by the scientific community and the information needed for decision making by user groups. This is supported by the fact that second-round survey responses and budgeting allocations placed a higher priority on *processes*, by which such needs can be reconciled, than on *products* themselves.

3. The workshop helped the participants to realize that given the difficulty – and in some cases, the impossibility – of reducing uncertainty in climate projections and achieving a perfect match between available and requested data, the ability to find compromises between desirable and actually available data is critical.

In other words, although in some cases, new model runs can produce the information required by user groups, quite often, the desired information cannot be provided. Nevertheless, other aspects of future climate changes with relevance for decision making can be assessed, which may still enable some progression of the decision process. Although iteration allows for better understanding, it does not necessarily result in changes in the needs of decision makers. We nevertheless suggest that IID can enable progress by reaching intermediate goals, allowing decision makers to articulate and revise their needs in conversation with researchers, and possibly to devise novel ways of producing proxy outputs.

After the workshop-seminar series, a student evaluation showed that student participants became proactive in incorporating interdisciplinary dialogue into their existing projects. Yet some kind of monitoring and longer term goals are necessary to ensure that the change of perspectives observed during the workshop persists. To this end, we suggest two mechanisms: applied case studies carried out over several months and modular integration into existing required courses within the curriculum (see Supporting Sections 7.7.4 and 7.7.5).

These findings suggest that concerted replication of similar events is a promising way to multiply the interactions between academics and decision makers, which are ultimately necessary to inform robust adaptation strategies. On the basis of our experience as organizers, the written evaluation of the workshop by the participants, and the oral feedback gathered, we identify three key recommendations for the organization of future events: i) create an interdisciplinary environment, ii) keep case studies manageable, and iii) provide concrete methods to deal with uncertainty (see Supporting Section 7.7.6).

In conclusion, we see such events as promising ways to intensify future interdisciplinary collaborations (Gornish et al., 2013) and produce competent facilitators to broker information between scientists and decision makers (Dilling and Lemos, 2011). Our experience shows that a workshop can reach, bring together, and benefit a wide range of participants, such as experts from industry, government, academia, humanitarian aid and development, and students and academic staff. As such, these workshops and classes are useful ways to complement and strengthen other channels fostering IID (e.g., boundary organizations and informal networks). We encourage the organization of similar events, with the midterm goal of improving adaptation strategies and better mitigating climate impacts.

Acknowledgements

We would like to thank all the speakers who contributed to our workshop: Niels Balzer, Nicole Clot, Thierry Corti, David Demeritt, Reto Knutti, Pete Fisher, Alan MacEachren, Daniel Maselli, and Rob Wilby. We would also like to pay tribute to the late Pete Fisher who we were very fortunate to have as an invited speaker at our workshop. David Demeritt proposed the participatory budgeting exercise described in the Supporting Section 7.7.3. We would also like to thank Silviya Nikolova for digitizing the survey cards, Ivo Heeb for assisting with coding the responses and the two anonymous reviewers whose comments helped to improve this commentary. We thank the Department of Geography, University of Zurich for funding through the Innovative Pool fund. Nans Addor attended the workshop 'Uncertainty in Climate Change: An Integrated Approach' in August 2012 at the National Center for Atmospheric Research in Boulder, USA from which inspiration for the workshop described in this commentary has been drawn. The data collected during the workshop are available upon request from the corresponding author.

7.7 Supporting information

7.7.1 Presentations

With presentations during the workshop, speakers introduced participants to the following topics:

1. Sources and types of uncertainties, including an uncertainty taxonomy (Fisher, 1999).
2. Uncertainty in climate models and scenarios (Deser et al., 2012a; Knutti and Sedláček, 2012; Knutti et al., 2013).
3. Visualizing uncertain information (MacEachren, 1992).
4. The social and political frameworks governing adaptation to future flooding in England (Demeritt et al., 2007).
5. Uncertainties in hydrological projections and their implications for water resource management, with a focus on developing more robust adaptation strategies (Wilby and Dessai, 2010).
6. Managing uncertainties and contradictory information in development cooperation (SDC Swiss Agency for Development and Cooperation, 2012).
7. Methods for handling uncertainty and climate risk in the reinsurance sector (Swiss Reinsurance, 2010).
8. The World Food Programme perspective on assessing uncertainty and risk relating to food security (Hazell et al., 2010; World Food Programme and Oxfam America, 2013).

9. Holistic assessment of vulnerabilities and adaptation measures in development cooperation (Clot, 2008; IISD, 2012).

7.7.2 Structure and evaluation of the opinion survey

To assess the influence of the workshop on participants' perspectives on exchanges between scientists and decision makers, we conducted an opinion survey before and after the workshop. Participants were asked *'What information and tools should be exchanged between researchers and decision makers to better address uncertainty?'*. Anonymous responses were solicited from all workshop participants (including the invited speakers) at the beginning of day 1 (Figure 7.3, round 1) and again at the conclusion of day 2 (round 2) using index cards with unique identifiers. To analyze patterns among responses and detect changes over the course of the workshop, responses were digitized and coded for content. Content analysis via coding is regularly employed in qualitative data analysis in social science and education research (Corbin and Strauss, 2008). Codes are researcher-generated constructs of a few words that distill concepts appearing in individual pieces of language-based data for the purpose of detecting patterns, creating categories, and building theory (Saldaña, 2013). In this case, coding was conducted by two collaborating individuals in two cycles using the qualitative data analysis software Atlas.ti. Initial or 'open' coding identified the discrete conceptual categories and relationships evidenced in the responses, followed by focused coding that grouped and refined initial codes based on their frequency and significance (Saldaña, 2013). From a total of 95 responses (53 pre-workshop and 42 post-workshop), we derived 57 codes. A simple frequency analysis allowed us to derive the code's relative frequency across the entirety of responses and to see whether this changed from round 1 to round 2. Codes with a relative frequency ≥ 0.10 in each round of responses were then selected for visualization.

7.7.3 Funding options for the mock budgetary exercise

For the mock budgetary exercise, each person was given three sticky notes marked '1', '2', and '3', and asked to rank their funding priorities. The proposals were the following and the participants' votes are summarized in Figure 7.4:

1. Invest in computational capacity and basic science to generate better climate models, and thus better predictions and understanding, which can then inform decision makers and other users.

2. Invest in instrumentation and data management to produce better, and more, observations.

3. Explore the potential of local knowledge and 'uncertified' expertise, prioritizing simplicity over sophistication.

4. Invest in better and more frequent communication, communication tools, and institutionalized forums; promote transparency of scientific assessments.

5. Facilitate dialogues at the science-practice interface to identify research questions and share findings.

6. Use the scientific risk assessment tools and principles of insurance to assess and manage risks in the face of uncertainty.

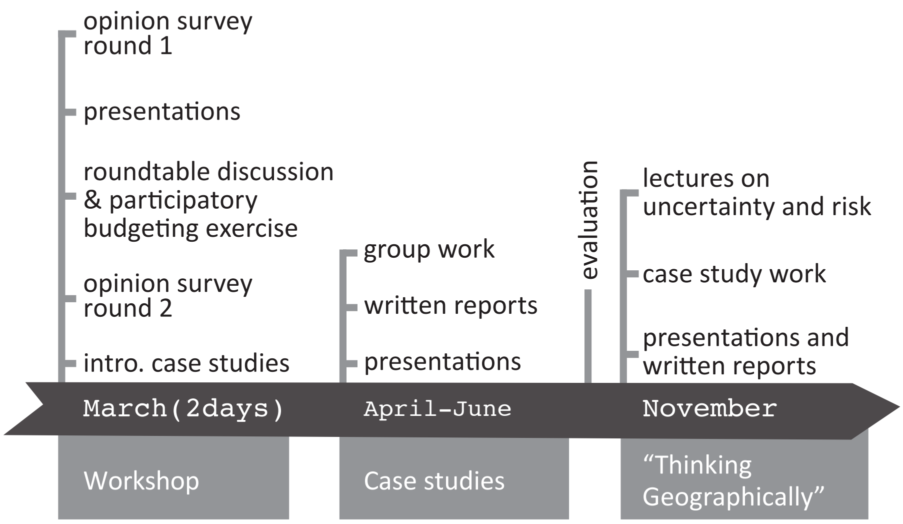

Figure 7.3: Overview of the components of the workshop, case study seminar and incorporation into a mandatory master's course for geography students entitled 'Thinking Geographically'.

7. Rather than focusing on uncertain future climate change impacts, invest in adaptation to extreme weather now.

8. Prioritize tools for sustainable development that reduce vulnerability through horizontal learning between communities rather than external technical solutions.

7.7.4 Case studies: exploring uncertainties in applied research

Following the workshop's expert presentations and roundtable discussion, students were broken into teams and assigned an extended real-world case study developed in collaboration with practitioners, many of whom were invited workshop speakers. The case studies required students from different subdisciplines of geography (geographic information science, remote sensing, physical geography, and human geography) to work in teams with members from bachelor to doctoral levels over a two-month period to identify and describe uncertainties for five applied cases, and further think about how these might influence or be incorporated into the decision-making process. Teams then presented their findings to each other in a one-day seminar.

The five case studies were developed based on both in-house expertise and on-going projects that would allow for an analysis of uncertainties by the student groups and facilitate discussion about the decision-making process. The case studies dealt with the uncertainties in:

1. a landslide early warning system in Colombia (Huggel et al., 2010),

2. the implemention of biochar technology in sugar cane farms in southern India (Singh et al., 2012),

3. flood forecasting and management for the city of Zurich (Demeritt et al., 2007; Addor et al., 2011), and

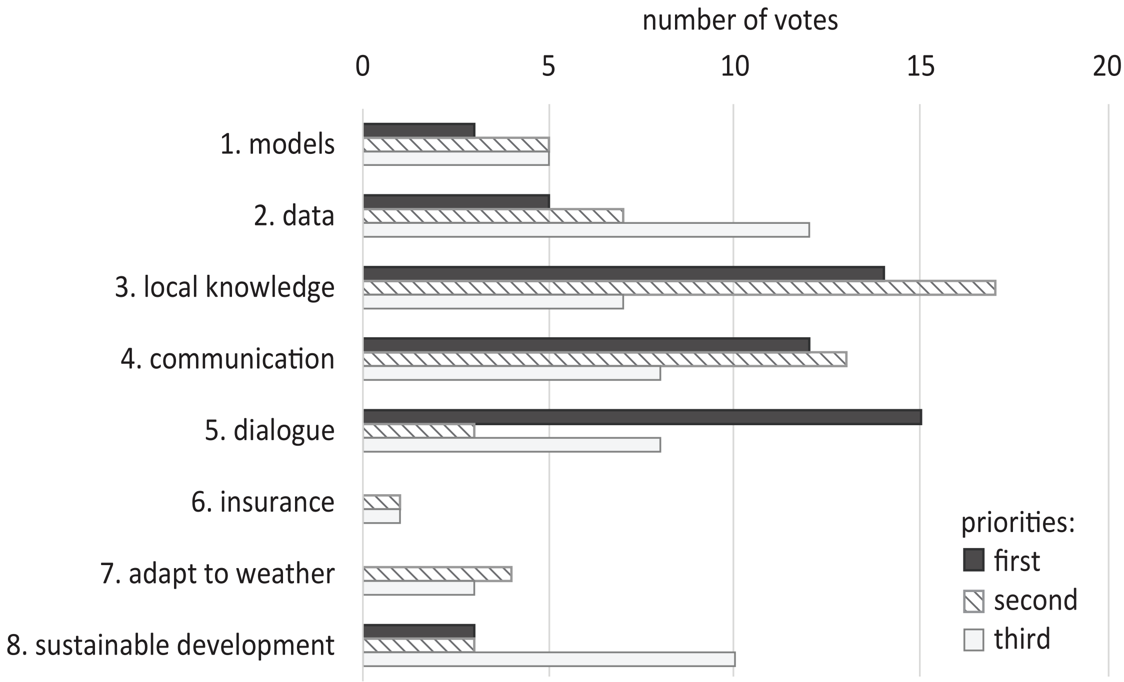

Figure 7.4: Summary of options to allocate hypothetical grant money, with participants' top three hypothetical budget priorities (with first priority votes indicated in solid dark gray, second priority votes with gray diagonal lines and third priority votes in solid light gray).

4. an adaptation and microinsurance project for food security in Ethiopia (World Food Programme and Oxfam America, 2013).

The fifth case study investigated the visualization of uncertainties for decision-making (MacEachren et al., 2005; Kleiner, 2013). Students looked critically at the visual depiction of uncertainties through a defined user experiment in order to understand how visualization choices influence a decision maker's perspective and the decision taken. The group developed an online survey that presented participants with two comparable visualizations, with one having no representation of uncertainty, and the other having some form of uncertainty representation. The results from this case study were incorporated into a master's thesis on this topic (Kleiner, 2013). Thinking about visualization and communication of results was an particularly useful exercise for students and one that could be developed into a more in-depth seminar or course module.

We found it important to provide the participants with concrete tools to deal with uncertainty. This is in particular supported by the following example. The case study that looked at implementing biochar technology in India came up with four different categories of uncertainties including those related to the production method (methodic uncertainties); the properties of biochar and its interaction with the agroecological system (environmental uncertainties); cost-benefit uncertainties (economic uncertainties); and uncertainties related to whether farmers would accept or 'like it' enough to actually implement the technology in their sugar cane farms (social uncertainties). Although identification of these uncertainties was a useful exercise, participants had difficulties thinking about what types of decision makers might use this information and how they might further be incorporated into decisions made.

7.7.5 'Thinking Geographically': incorporation of workshop elements into the curriculum

Building on the workshop case studies and feedback, we incorporated some of the material into an already-existing course within the department's master's curriculum called 'Thinking Geographically'. Three short lectures drew on material from the workshop, providing perspectives from physical, social, and geographic information sciences as well as top-down versus bottom-up approaches for assessing impacts and approaching adaptation. Students in the course then went on to work on case studies again in teams, with a focus this time on broader topics. In comparison with the case studies linked to the workshop, these less narrowly defined cases encouraged more exploratory discussion within the groups about risks and potential uncertainties within decision-making processes, especially related to more qualitative aspects stemming from the human geography perspective, and more critical reflection on the methodology applied for risk assessment and management.

7.7.6 What are the key elements for productive dialogue-generating academic events?

1. Create an interdisciplinary environment. Feedback for the two-day workshop was very positive, with two thirds of the participants rating it as 'very successful' in fostering cross-disciplinary learning and discussion. The speakers commented among themselves on how few occasions they normally have to exchange on their experiences on dealing with uncertainties. The students also found that the workshop's different format from ordinary lectures encouraged them to talk freely and raise questions, but thought the two-day format was too brief. Three months after the workshop, many physical geography students reported that they had found new motivation to integrate human geography and development studies into their curriculum. Professors attending the workshop also found new inspiration, and incorporated elements of the workshop into their classes.

2. Keep case studies manageable. Although the students praised the multi-level teams of bachelor, master, and PhD groups formed to work on case studies, two thirds rated them only 'somewhat successful'. This average evaluation likely reflects the fact that the participants were given quite difficult tasks related to their case studies in comparison to the more positively evaluated workshop, in which they listened to engaging presentations from speakers. The difficulties they had with the case study exercises nevertheless congruously emphasized the real-world difficulty in addressing such challenging and interdisciplinary issues. Indeed, the difficulty in carrying out the tasks in the case studies also reflects the barriers to introducing interdisciplinary climate change issues into curriculum (Davison et al., 2014).

3. Provide concrete methods to deal with uncertainty. Although participants found case studies to be a useful exercise, many wished for more concrete decision making scenarios and analytical tools that could have been used and applied within the case studies (see the example in Supporting Section 7.7.4). Overall, if the workshop were to be organized again, our main recommendation would be to familiarize participants with concrete technical methods to facilitate decision making in the face of uncertainty. For instance, for the case studies, the participants could establish, possibly in a simplified way, the vulnerabilities of a community or infrastructure, and would

then be provided with some uncertain climate information that they would assess and use to explore future risks (see e.g., the visuals in Brown and Wilby, 2012) and make recommendations. They could then be given a complementary set of climate data, such as simulations from the latest climate models, to introduce the importance of flexible adaptation frameworks.

Chapter 8

Conclusions and outlook

Note: The conclusions are encapsulated in three sections, each of them being followed by a related outlook section, whose title is a question.

8.1 The mismatch between RCM simulations and observations

Discharge projections are based on the combination of climate models and hydrological models. This combination is complicated by systematic differences between climate model outputs and observations. Pragmatic methods exist to coerce model simulations so that they look like observations, but their ability to correct simulations under future conditions is questionable, in particular because they do not address nor identify the reasons behind this mismatch. In this thesis, we investigated one reason behind the mismatch between model outputs and observations: errors in the synoptic atmospheric circulation, which primarily stem from the general circulation model (GCM) driving the regional climate model (RCM). We used circulation types to characterize the atmospheric circulation and to address the research questions outlined in Section 1.7 for Chapter 4. We found that GCM-RCMs capture the frequency and regime of circulation types under present climate reasonably well, but that some severe circulation biases exist. We reported in particular a systematic overestimation of the frequency of westerly flow in winter and we demonstrated its contribution to the overestimation of winter precipitation over Switzerland. Precipitation and temperature simulations are overall more affected by circulation biases in winter than in summer. We investigated how well one of the most widely used methods for bias-correction (quantile mapping) deals with theses biases. We found that circulation biases lead to residual errors in the 'bias-corrected' time series. We demonstrated that the amplitude of these residual errors can be large in case of large circulation biases, and may deteriorate impact simulations (Chapter 4). The shortcomings of quantile mapping also clearly appeared when investigating variables important for the modelling of rare and damaging hydrological events, such as floods and droughts. We found that the quantile mapping method falls short. It corrects well for biases in daily variables, such as the frequency of wet days, but fails to capture multi-day statistics relevant for extreme events, such as precipitation over 4 consecutive days (Chapter 5). Our results stress the importance of better identifying and accounting for the origins of biases in RCM simulations.

Although the mismatch between RCM simulations and observations is usually referred to as 'model bias', suggesting that it is caused by model limitations alone, we investigated the contribution to this mismatch of interpolation errors in observational datasets and natural climate variability. Interpolation errors are an avoidable component of RCM-observation comparisons, because RCMs are run on a grid and observations are collected by stations on an irregular network. Hence, either observations are upscaled to the RCM grid, or RCM are downscaled to station locations. In either case errors are introduced. Natural variability relates to the chaotic nature of the climatic system: even a perfect model will depart from observations because it can be in a different (e.g., drier or warmer) state than the observed system. We found that interpolation errors contribute to a major part of the RCM-observation differences, in particular in areas of complex topography. This implies that different datasets will lead to different biases (and performance of the model), and hence that bias-correction should account for observational uncertainty, for instance by using several observational datasets. As for the research questions raised in Section 1.7 for Chapter 3, we found that the influence of natural variability on the results of climate model evaluation over multi-decadal periods was comparatively small. Hence for 30-year means of temperature and precipitation, on which our study focussed, we argue that the unpredictable nature of internal variability does not significantly preclude our chances to reduce climate model biases by post-processing. The contribution of the natural variability might nevertheless be significant for other variables, for instance extremes or trends. In this case, we propose that both model evaluation and bias-correction should rely on several model runs reflecting natural variability. We conclude that a sensible evaluation of climate models and interpretation of their simulations requires a differentiation between model errors, uncertainties in observational datasets and natural variability (Chapter 3).

8.2 How sensitive are discharge projections to errors in climate inputs?

So far the effects of bias-correction on the simulated discharge have only been investigated by a few studies (Cloke et al., 2012; Muerth et al., 2013; Teng et al., 2015). It is for example still unclear how uncertainties in regional observations affect the bias-correction of climate simulations and the resulting projected discharge. But ensembles of observations are on the rise. They already exist at the global scale (Morice et al., 2012), and ensembles of observations at the regional scale, for instance for Switzerland (C. Frei, 2014, personal communication) and USA (E. Gutmann, 2014, personal communication) are on their way. Using them for bias-correction and hydrological projections should provide new insights into the contribution of observational uncertainties to the uncertainty in discharge projections, a topic little explored so far.

Errors in climate simulations can lead to delayed errors in discharge simulations. For instance, the overestimation of winter precipitation discussed in this thesis might have implications later in the year in RCM and hydrological simulations. The snow pack and especially the soil are large water reservoirs, with a long-lasting memory. Too much rain in winter can lead to an excess of soil moisture for months, which may bias the energy balance (more latent heat, less sensible heat) possibly leading to an underestimation of spring temperature. If this can be observed, it would be another example of error propagation in model chains.

Further, it is expected that the simulation of different discharge variables (e.g., mean winter discharge or flood frequency) will be sensitive to different kinds of error in the climate inputs, but the exact links still have to be uncovered. For instance, as stressed in Chapter 4, hydrological models might be little sensitive to residual errors in the bias-corrected RCM simulations if those errors are smaller than observation errors. We showed in Chapter 2 that the two post-processing techniques led to similar changes in the hydrological regime, but they might deliver different projections if extremes were considered instead. Further, whether these errors matter for adaptation strategies will depend on the question being addressed (Chapter 7).

This thesis focussed on changes in mean temperature and precipitation and in the hydrological regime, and not on extremes. This choice was made for two reasons. Firstly, the correct simulation of mean quantities is a pre-requirement for the correct simulation of extremes, for instance because mean quantities strongly influence soil moisture, which has a crucial influence on the occurrence or not of a flood for a severe precipitation event. Second, although many impacts are related to punctual extreme events, many impacts are driven by the baseline, for instance the input of water into a reservoir. That said, several research questions addressed in this thesis could be reformulated to focus on extremes instead of mean quantities.

8.3 The contribution of hydrological models to projection uncertainty

In Chapter 2, we used a particularly large ensemble of model chains to better understand which sources of uncertainty affect discharge projections most, and to assess the influence of hydrological models on the simulated future discharge. This setup allowed us to answer the research questions defined in Section 1.7 for Chapter 2. We found that the climate models and natural climate variability have overall the most significant contribution to the variance of the projected discharge, but the choice of emission scenario plays a large role by the end of the 21st century. Further, the contribution of the hydrological models to the projection uncertainty varies strongly with catchment elevation.

We found that in lower elevation catchments, the contribution of the hydrological models to the overall uncertainty was barely significant. In other words, both simple and more complex models delivered similar projections. We proposed that in this situation, future discharge changes such as the earlier snowmelt peak and lower summer discharge are primarily driven by the projected increase of temperature and decrease of summer precipitation, to an extent that the differences between the models are overruled. This led to the emergence of robust changes, despite uncertainties in the hydrological projections. It should however not be concluded that hydrological model complexity does not matter for discharge projections, but rather that uncertainty needs to be put in its context, related to the uncertainty and robustness of the other elements of the model chain. In higher elevation catchments, where snow, glaciers, dams and topography make hydrological modeling challenging, significant differences appeared between the simulations of the different hydrological models, making them a significant source of uncertainty. This points towards deficiencies in the representation and calibration of atmospheric and hydrological processes in catchments of complex topography.

In Chapter 6 we investigated trends in the water balance of four study catchments over 1971-2010. Preliminary results indicate that although temperature increased significantly, there were no significant trends in actual evapotranspiration estimates. Trends in precipitation records were not significant. The ability of the conceptual model HBV to reproduce observed trends in discharge and evapotranspiration with constant parameter values was overall good, which suggests more robustness that might have been expected. When HBV was calibrated over 10-year periods, trends in parameter values appeared, but we could not associate them with certainty to changes in hydrological processes. In fact, we adopted a 'trading time for space' approach, which revealed that some trends in space were inconsistent with trends in time, and underscored the challenge of physically justifying the transfer of parameter values in time and space. Ongoing simulations will help to clarify the ability of HBV to simulate the water balance in changing conditions and to further test the rationale behind time-dependent parameter values.

8.4 Which hydrological models for impact studies?

Hydrological models used in impact studies were usually not explicitly designed to be run under changing conditions. Instead, the community generally uses models that perform well under present climate, but instead of forcing them with observations, forces them with climate projections. This leads to the question: How to design a hydrological model well-suited for climate change impact studies? A way to address this question would be to run competing models in parallel and compare their performance. Such an approach based on an ensemble would however be confronted with four main challenges:

1. The effect of the different components of the hydrological models are confounded. It is not necessarily clear why a particular model performs well. Is it for instance because of the way its parameters were estimated, its spatial discretization or its representation of ground water? Similarly, the good performance of some model components can be masked by the poor performance of other components of the same model.

2. Ensembles are not explicitly designed to sample model uncertainty. Instead, models are typically developed by different institutes, working toward the 'best model', but usually not making a concerted effort to chose different model structures or process representations in order to sample model uncertainty. Models hence lack independence and exhibit similar biases as a result of their similar assumptions.

3. Different execution and calibration modes make increasing the number of models difficult. Adding new models to the ensemble helps to sample model uncertainty, but because each model is designed and run separately (they for instance need different input files and run under different environments), adding new models is usually cumbersome.

4. Model performance under future conditions is difficult to assess. Models performing well under present conditions do not necessarily capture well changes induced by global warming (Racherla et al., 2012). Further, model performance can significantly decrease if the model is run in climatic conditions significantly different from those of its calibration period (Merz et al., 2011; Coron et al., 2012; Seiller et al., 2012; Gharari et al., 2013).

To address these challenges, an option is to use the newly developed Structure for Unifying Multiple Modeling Alternatives (SUMMA, Clark et al., 2015a,b). SUMMA is a modular framework allowing the modeler to chose from a range of i) spatial architectures and discretization, ii) formulations of individual biophysical and hydrologic processes, iii) numerical methods to solve model equations. A key strength of SUMMA is that competing formulations can be compared by being successively used in the model, keeping the rest of the model constant. This allows for the separation of their effect from that of the rest of the model, and thereby for the systematic evaluation of model structure and formulation (this addresses challenge 1). SUMMA enables the systematic evaluation and comparison of a wide range of competing model formulations, thereby yielding a vastly improved sampling of simulation uncertainty (challenge 2). Using a unique structure allows for the recycling of input files (issue 3). Further, models should not only be evaluated based on their ability to reproduce observed events, the stability of their performance under changing conditions should be scrutinized as well for instance using differential split-sampling tests (Klemeš, 1986; Andréassian et al., 2009; Refsgaard et al., 2013). This addresses challenge 4. Finally, as discussed in Chapter 2, in order to generalize results and assess their robustness, hydrological modeling should to be conducted for a wide range of catchments (Gupta et al., 2014).

SUMMA hence enables to explore hydrological model structures and formulations in a more systematic way and for a wider range of catchments than studies thus far. This would allow to test how strongly structure and formulation requirements are related to catchment characteristics, systematically explore the relationship between model complexity and model performance, and finally, explore the sensitivity of discharge projections to competing model structures.

8.5 Toward stronger synergies between research and decision-making

We found that significant efforts to reduce of greenhouse gas emissions (RCP2.6) lead to a decrease of the projected impacts on hydrological regimes by about a factor of two (Chapter 2). This might stimulate the adoption of policies aiming toward the reduction our carbon footprint. Further, in terms of adaptation, we considered important to identify robust changes, since if decisions had to be made based on the projections currently available, we would recommend to base them on the changes considered most robust. We showed that although irreducible sources of uncertainty exist (natural variability), robust changes can be isolated. We consider these results relevant and they were part of the CH2014-Impacts report (Rössler et al., 2014), but it is unclear whether they will influence policy or be used for decision-making. Have they reached decision-makers and if yes, did they provide them the information they needed?

A lot remains to be done to strengthen links between climate researchers and decision-makers, and a key challenge is to better handle uncertainties, which are inherent to climate change. We argue that a way to proceed is to combine the dominant ways to deal with uncertainty in decision-making processes (bottom-up and top-down approaches), and do so through iterative, open-ended interdisciplinary dialogue. Just like advancing hydrological research relies on the dialogue between field hydrologists and modelers (Seibert and McDonnell, 2002), we propose that strengthening interactions and mutual understanding between the climate community and end-users can help adaptation to progress. In Chapter 7 we showed that targeted events, such as the workshop we organized, can generate a new basis for this dialogue. This workshop provided the participants with an overview of the current research on uncertainty and on how uncertainty is dealt with by decision-makers, it brought together a wide range of scientists, practitioners and students, and it exposed students at an early stage of their professional life to multidisciplinary collaborations. To answer the research questions outlined in Section 1.7 for Chapter 7, we conducted a survey before and after the workshop, which revealed a shift in perspectives about the relationship between researchers and user groups. Before the workshop, participants considered that to better address uncertainty, new products had to be generated by academic researchers based on their expertise and their apparently more direct access to climate information. In contrast, participants' responses after the event tended to prioritize processes rather than products. For instance, previously prioritized outputs such as visualizations were virtually discounted, and recommendations to produce quantifications of uncertainty and metadata were only half as frequent as before the event. Instead, the most common recommendations were for dialogue and more frequent and improved channels for communication. We propose that such academic events can contribute to foster iterative and interdisciplinary exchanges, with the midterm goal of improving adaptation strategies and better mitigating climate impacts.

8.6 Case studies to close the loop between research and decision-making?

In the weather analogy used in the introduction of this thesis, we stressed the importance of producing results tailored to end-users' needs, which actually help them with decision-making. Identifying their needs requires a direct dialogue. For instance, in the case of the management of a reservoir, it is essential to sit down with its operators, find out under which circumstances their activities are at risk, but also, understand how well we can model these conditions in the future. Those two aspects are arguably very activity- and site-specific, which suggests that studies addressing these questions might in fact be case studies.

Such studies will probably include more sources of uncertainty than investigated in this thesis. For instance those related to the cost of a flood, to future energetic needs or even to the perception of risk. Dialogue should help to determine how to best include these uncertainties into the decision-making process. From an other perspective, such case studies have the potential to help us to better understand the sensitivity of end-users to imperfections in the model chain. It might not be necessary to reduce these uncertainties, but rather, to characterize them in a reliable way. Therefore, a key element is to understand where uncertainties are generated in the model chain and how they propagate down the chain. This thesis provided some insights into this.

8.7 Concluding remarks on models in climate impact studies

Some might argue that in this thesis, the focus on model failure is too strong. But as Beven (2007) puts it, 'more may be learned from model rejection than acceptance; rejection of a hypothesis, when properly justified, is an important stage in model development and improvement'. It is hence crucial to assess what models can and cannot do, when they shine and when they fail. By reporting and exploring the reasons behind their limitations, we recognize that models are simplifications of reality (Andréassian et al., 2010). Further, even though we report several cases of model failures, they do not prevent us from finding robust changes in discharge projections.

To conclude, let us go back to the quote that opened this thesis: '[...] models' usefulness rests in their ability to help us think about a system - as heuristic tools to help us understand what we can actually observe or estimate, or as a way to challenge existing formulations and intuitions' (Weaver et al., 2013). When models are used for impact modeling, a lot of the attention typically goes to their outputs. It is important to know how discharge will change in the future, yet it might be even more important to explore how robust and uncertain projections are, and how to use them for decision-making. In conclusion, exploring abilities and limitations of models, and engaging in interdisciplinary dialogue, is essential to benefit from the full potential of models in the context of climate change.

Bibliography

Adam, J. C., Hamlet, A. F., and Lettenmaier, D. P.: Implications of global climate change for snowmelt hydrology in the twenty-first century, Hydrological Processes, 23, 962–972, doi:10.1002/hyp.7201, 2009.

Addor, N. and Fischer, E. M.: The influence of natural variability and interpolation errors on bias characterization in RCM simulations, Journal of Geophysical Research: Atmospheres, 120, 180–10,195, doi:10.1002/2014JD022824, 2015.

Addor, N. and Seibert, J.: Bias-correction for hydrological impact studies - beyond the daily perspective, Hydrological Processes, 28, 4823–4828, doi:10.1002/hyp.10238, 2014.

Addor, N., Jaun, S., Fundel, F., and Zappa, M.: An operational hydrological ensemble prediction system for the city of Zurich (Switzerland): skill, case studies and scenarios, Hydrology and Earth System Sciences, 15, 2327–2347, doi:10.5194/hess-15-2327-2011, 2011.

Addor, N., Rössler, O., Köplin, N., Huss, M., Weingartner, R., and Seibert, J.: Robust changes and sources of uncertainty in the projected hydrological regimes of Swiss catchments, Water Resources Research, 50, 7541–7562, doi:10.1002/2014WR015549, 2014.

Addor, N., Ewen, T., Johnson, L., Çöltekin, A., Derungs, C., and Muccione, V.: From products to processes: Academic events to foster interdisciplinary and iterative dialogue in a changing climate, Earth's Future, 3, 289–297, doi:10.1002/2015EF000303, 2015.

Andréassian, V., Perrin, C., Berthet, L., Le Moine, N., Lerat, J., Loumagne, C., Oudin, L., Mathevet, T., Ramos, M.-H., and Valéry, A.: HESS Opinions "Crash tests for a standardized evaluation of hydrological models", Hydrology and Earth System Sciences, 6, 3669–3685, doi:10.5194/hessd-6-3669-2009, 2009.

Andréassian, V., Perrin, C., Parent, E., and Bárdossy, A.: The Court of Miracles of Hydrology: can failure stories contribute to hydrological science?, Hydrological Sciences Journal, 55, 849–856, doi:10.1080/02626667.2010.506050, 2010.

Ban, N., Schmidli, J., and Schär, C.: Evaluation of the convection-resolving regional climate modeling approach in decade-long simulations, Journal of Geophysical Research: Atmospheres, 119, 7889–7907, doi:10.1002/2014JD021478, 2014.

Bárdossy, A. and Pegram, G.: Downscaling precipitation using regional climate models and circulation patterns toward hydrology, Water Resources Research, 47, 1–18, doi:10.1029/2010WR009689, 2011.

Barnett, T. P., Adam, J. C., and Lettenmaier, D. P.: Potential impacts of a warming climate on water availability in snow-dominated regions., Nature, 438, 303–9, doi:10.1038/nature04141, 2005.

Begert, M., Schlegel, T., and Kirchhofer, W.: Homogeneous temperature and precipitation series of Switzerland from 1864 to 2000, International Journal of Climatology, 25, 65–80, doi:10.1002/joc.1118, 2005.

Bellprat, O., Kotlarski, S., Lüthi, D., and Schär, C.: Exploring Perturbed Physics Ensembles in a Regional Climate Model, Journal of Climate, 25, 4582–4599, doi:10.1175/JCLI-D-11-00275.1, 2012.

Bellprat, O., Kotlarski, S., Lüthi, D., and Schär, C.: Physical constraints for temperature biases in climate models, Geophysical Research Letters, 40, 4042–4047, doi:10.1002/grl.50737, 2013.

Berghuijs, W. R., Woods, R. A., and Hrachowitz, M.: A precipitation shift from snow towards rain leads to a decrease in streamflow, Nature Climate Change, pp. 18–21, doi:10.1038/NCLIMATE2246, 2014.

Beven, K.: Changing ideas in hydrology - The case of physically based models, Journal of Hydrology, 105, 157–172, doi:10.1016/0022-1694(90)90161-P, 1989.

Beven, K.: A manifesto for the equifinality thesis, Journal of Hydrology, 320, 18–36, doi:10.1016/j.jhydrol.2005.07.007, 2006.

Beven, K.: Environmental Modelling: An Uncertain Future?, Routledge, London and New York, URL http://www.uncertain-future.org.uk, 2008.

Beven, K. and Young, P.: A guide to good practice in modeling semantics for authors and referees, Water Resources Research, 49, doi:10.1002/wrcr.20393, 2013.

Beven, K. J.: Working towards integrated environmental models of everywhere : uncertainty, data, and modelling as a learning process, Hydrology and Earth System Sciences, 11, 460–467, 2007.

Beven, K. J. and Cloke, H. L.: Comment on "Hyperresolution global land surface modeling: Meeting a grand challenge for monitoring Earth's terrestrial water," by Eric F. Wood et al., Water Resources Research, 48, W01 801, doi:10.1029/2011WR010982, 2012.

Bezzola, G. R. and Hegg, C., eds.: Ereignisanalyse Hochwasser 2005, Teil 1 - Prozesse, Schäden und erste Einordnung, Bundesamt für Umwelt BAFU, Eidgenössische Forschungsanstalt WSL, Switzerland, 2007.

Bidwell, D., Dietz, T., and Scavia, D.: Fostering knowledge networks for climate adaptation, Nature Climate Change, 3, 610–611, doi:10.1038/nclimate1931, 2013.

Bierkens, M. F. P., Bell, V. A., Burek, P., Chaney, N., Condon, L. E., David, C. H., de Roo, A., Döll, P., Drost, N., Famiglietti, J. S., Flörke, M., Gochis, D. J., Houser, P., Hut, R., Keune, J., Kollet, S., Maxwell, R. M., Reager, J. T., Samaniego, L., Sudicky, E., Sutanudjaja, E. H., van de Giesen, N., Winsemius, H., and Wood, E. F.: Hyperresolution global hydrological modelling: what is next?, Hydrological Processes, 29, 310–320, doi:10.1002/hyp.10391, 2015.

Birsan, M.-V., Molnar, P., Burlando, P., and Pfaundler, M.: Streamflow trends in Switzerland, Journal of Hydrology, 314, 312–329, doi:10.1016/j.jhydrol.2005.06.008, 2005.

Blandford, T. R., Humes, K. S., Harshburger, B. J., Moore, B. C., Walden, V. P., and Ye, H.: Seasonal and Synoptic Variations in Near-Surface Air Temperature Lapse Rates in a Mountainous Basin, Journal of Applied Meteorology and Climatology, 47, 249–261, doi:10.1175/2007JAMC1565.1, 2008.

Blenkinsop, S., Jones, P. D., Dorling, S. R., and Osborn, T. J.: Observed and modelled influence of atmospheric circulation on central England temperature extremes, International Journal of Climatology, 1660, 1642–1660, doi:10.1002/joc, 2009.

Blöschl, G. and Montanari, A.: Climate change impacts - throwing the dice?, Hydrological Processes, 381, 374–381, doi:10.1002/hyp.7574, 2010.

Boberg, F. and Christensen, J. H.: Overestimation of Mediterranean summer temperature projections due to model deficiencies, Nature Climate Change, 2, 433–436, doi:10.1038/nclimate1454, 2012.

Bordoy, R. and Burlando, P.: Stochastic downscaling of climate model precipitation outputs in orographically complex regions: 2. Downscaling methodology, Water Resources Research, 50, doi:10.1002/wrcr.20443, 2014.

Bosshard, T., Kotlarski, S., Ewen, T., and Schär, C.: Spectral representation of the annual cycle in the climate change signal, Hydrology and Earth System Sciences, 15, 2777–2788, doi:10.5194/hess-15-2777-2011, 2011.

Bosshard, T., Carambia, M., Goergen, K., Kotlarski, S., Krahe, P., Zappa, M., and Schär, C.: Quantifying uncertainty sources in an ensemble of hydrological climate-impact projections, Water Resources Research, 49, 1523–1536, doi:10.1029/2011WR011533, 2013a.

Bosshard, T., Kotlarski, S., Zappa, M., and Schär, C.: Hydrological climate-impact projections for the Rhine river: GCM-RCM uncertainty and separate temperature and precipitation effects, Journal of Hydrometeorology, p. 131121143357000, doi:10.1175/JHM-D-12-098.1, 2013b.

Brown, C. and Wilby, R.: An alternate approach to assessing climate risks, Eos, Transactions American Geophysical Union, 92, 92–94, doi:10.1029/2012EO410001, 2012.

Brown, C., Werick, W., Leger, W., and Fay, D.: A Decision-Analytic Approach to Managing Climate Risks: Application to the Upper Great Lakes, Journal of the American Water Resources Association, 47, 524–534, doi:10.1111/j.1752-1688.2011.00552.x, 2011.

Brown, C., Ghile, Y., Laverty, M., and Li, K.: Decision scaling: Linking bottom-up vulnerability analysis with climate projections in the water sector, Water Resources Research, 48, doi:10.1029/2011WR011212, 2012.

Bundesamt für Umwelt BAFU (Eds.): Auswirkungen der Klimänderung auf Wasserressourcen und Gewässer: Synthesebericht zum Projekt "Klimaveränderung und Hydrologie in der Schweiz" (CCHydro), Bundesamt für Umwelt, Bern, Switzerland, 2012.

Buser, C. M., Künsch, H. R., Lüthi, D., Wild, M., and Schär, C.: Bayesian multi-model projection of climate: bias assumptions and interannual variability, Climate Dynamics, 33, 849–868, doi:10.1007/s00382-009-0588-6, 2009.

Cash, D. W., Borck, J. C., and Patt, A. G.: Countering the Loading-Dock Approach to Linking Science and Decision Making: Comparative Analysis of El Nino/Southern Oscillation (ENSO) Forecasting Systems, Science, Technology & Human Values, 31, 465–494, doi:10.1177/0162243906287547, 2006.

Cattiaux, J., Douville, H., and Peings, Y.: European temperatures in CMIP5: Origins of present-day biases and future uncertainties, Climate Dynamics, 41, 2889–2907, doi:10.1007/s00382-013-1731-y, 2013.

Ceppi, P., Scherrer, S. C., Fischer, A. M., and Appenzeller, C.: Revisiting Swiss temperature trends 1959-2008, International Journal of Climatology, doi:10.1002/joc.2260, 2010.

CH2011: Swiss Climate Change Scenarios CH2011, C2SM, MeteoSwiss, ETH, NCCR Climate, and OcCC, Zurich, Switzerland, 2011.

CH2014 Impacts: Toward quantitative scenarios of climate change impacts in Switzerland, FOEN, MeteoSwiss, C2SM, Agroscope, and ProClim, Bern, Switzerland, 2014.

Chen, C., Haerter, J. O., Hagemann, S., and Piani, C.: On the contribution of statistical bias correction to the uncertainty in the projected hydrological cycle, Geophysical Research Letters, 38, 1–6, doi:10.1029/2011GL049318, 2011.

Chen, J., Brissette, F. P., Chaumont, D., and Braun, M.: Finding appropriate bias correction methods in downscaling precipitation for hydrologic impact studies over North America, Water Resources Research, 49, 4187–4205, doi:10.1002/wrcr.20331, 2013.

Chen, J., Brissette, F. P., and Lucas-Picher, P.: Assessing the limits of bias-correcting climate model outputs for climate change impact studies, Journal of Geophysical Research: Atmospheres, 120, 1123–1136, doi:10.1002/2014JD022635.Received, 2015.

Christensen, J. H. and Christensen, O. B. s.: A summary of the PRUDENCE model projections of changes in European climate by the end of this century, Climatic Change, 81, 7–30, doi:10.1007/s10584-006-9210-7, 2007.

Christensen, J. H., Boberg, F., Christensen, O. B., and Lucas-Picher, P.: On the need for bias correction of regional climate change projections of temperature and precipitation, Geophysical Research Letters, 35, doi:10.1029/2008GL035694, 2008.

Clark, M. P., Slater, A. G., Rupp, D. E., Woods, R. A., Vrugt, J. A., Gupta, H. V., Wagener, T., and Hay, L. E.: Framework for Understanding Structural Errors (FUSE): A modular framework to diagnose differences between hydrological models, Water Resources Research, 44, doi:10.1029/2007WR006735, 2008.

Clark, M. P., Nijssen, B., Lundquist, J. D., Kavetski, D., Rupp, D. E., Woods, R. A., Freer, J. E., Gutmann, E. D., Wood, A. W., Brekke, L. D., Arnold, J. R., Gochis, D. J., and Rasmussen, R. M.: A unified approach for process-based hydrologicmodeling: 1. Modeling concept, Water Resources Research, 51, 2498–2514, doi:10.1002/2015WR017198, 2015a.

Clark, M. P., Nijssen, B., Lundquist, J. D., Kavetski, D., Rupp, D. E., Woods, R. A., Freer, J. E., Gutmann, E. D., Wood, A. W., Gochis, D. J., Rasmussen, R. M., Tarboton,

D. G., Mahat, V., Flerchinger, G. N., and Marks, D. G.: A unified approach for process-based hydrologic modeling: 2. Model implementation and case studies, Water Resources Research, 51, 2515–2542, doi:10.1002/2015WR017200, 2015b.

Cloke, H. L., Wetterhall, F., He, Y., Freer, J. E., and Pappenberger, F.: Modelling climate impact on floods with ensemble climate projections, Quarterly Journal of the Royal Meteorological Society, doi:10.1002/qj.1998, 2012.

Clot, N.: Highlighting local coping strategies for drought, in: Integrating Disaster Risk Reduction into Climate Change Adaptation, UN/ISDR Gender Perspective, 2008.

Collins, H. M. and Evans, R.: The Third Wave of Science Studies: Studies of Expertise and Experience, Social Studies of Science, 32, 235–296, 2002.

Corbin, J. and Strauss, A.: Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory, Sage Publications, 2008.

Coron, L., Andréassian, V., Perrin, C., Lerat, J., Vaze, J., Bourqui, M., and Hendrickx, F.: Crash testing hydrological models in contrasted climate conditions: An experiment on 216 Australian catchments, Water Resources Research, 48, doi:10.1029/2011WR011721, 2012.

Coron, L., Andréassian, V., Perrin, C., Bourqui, M., and Hendrickx, F.: On the lack of robustness of hydrologic models regarding water balance simulation: a diagnostic approach applied to three models of increasing complexity on 20 mountainous catchments, Hydrology and Earth System Sciences, 18, 727–746, doi:10.5194/hess-18-727-2014, 2014.

Cortés, G., Vargas, X., and McPhee, J.: Climatic sensitivity of streamflow timing in the extratropical western Andes Cordillera, Journal of Hydrology, 405, 93–109, doi:10.1016/j.jhydrol.2011.05.013, 2011.

Court, A.: Measures of streamflow timing, Journal of Geophysical Research, 67, 4335–4339, doi:10.1029/JZ067i011p04335, 1962.

Crow, D. and Boykoff, M.: Introduction, in: Culture, Politics and Climate Change: How information shapes our common future, edited by Crow, D. and Boykoff, M., Routledge, New York, 2014.

Dabelko, G. D.: Speaking Their Language: How to Communicate Better with Policy-makers and Opinion Shapers - and Why Academics Should Bother in the First Place, International Environmental Agreements: Politics, Law and Economics, 5, 381–386, doi:10.1007/s10784-005-8329-8, 2005.

Davison, A., Brown, P., Pharo, E., Warr, K., McGregor, H., Terkes, S., Boyd, D., and Abuodha, P.: Distributed leadership: Building capacity for interdisciplinary climate change teaching at four universities, International Journal of Sustainability in Higher Education, 15, 98–110, doi:10.1108/IJSHE-10-2012-0091, 2014.

Demeritt, D., Cloke, H., Pappenberger, F., Thielen, J., Bartholmes, J., and Ramos, M.: Ensemble predictions and perceptions of risk, uncertainty, and error in flood forecasting, Environmental Hazards, 7, 115–127, doi:10.1016/j.envhaz.2007.05.001, 2007.

Deser, C., Knutti, R., Solomon, S., and Phillips, A. S.: Communication of the role of natural variability in future North American climate, Nature Climate Change, 2, 775–779, doi:10.1038/nclimate1562, 2012a.

Deser, C., Phillips, A., Bourdette, V., and Teng, H.: Uncertainty in climate change projections: the role of internal variability, Climate Dynamics, 38, 527–546, doi:10.1007/s00382-010-0977-x, 2012b.

Dessai, S. and Hulme, M.: Does climate adaptation policy need probabilities?, Climate Policy, 4, 107–128, doi:10.3763/cpol.2004.0411, 2004.

Dessai, S., Hulme, M., Lempert, R., and Pielke, R.: Do We Need Better Predictions to Adapt to a Changing Climate?, Eos, Transactions American Geophysical Union, 90, 111, doi:10.1029/2009EO130003, 2009.

Dilling, L. and Lemos, M. C.: Creating usable science: Opportunities and constraints for climate knowledge use and their implications for science policy, Global Environmental Change, 21, 680–689, doi:10.1016/j.gloenvcha.2010.11.006, 2011.

Dobler, C., Hagemann, S., Wilby, R. L., and Stötter, J.: Quantifying different sources of uncertainty in hydrological projections in an Alpine watershed, Hydrology and Earth System Sciences, 16, 4343–4360, doi:10.5194/hess-16-4343-2012, 2012.

Doherty, J.: PEST: Software for Model-Independent Parameter Estimation., Tech. rep., Watermark Numerical Computing, Australia., URL `http://www.sspa.com/pest`, 2005.

Dosio, A. and Paruolo, P.: Bias correction of the ENSEMBLES high-resolution climate change projections for use by impact models: Evaluation on the present climate, Journal of Geophysical Research, 116, 1–22, doi:10.1029/2011JD015934, 2011.

Ehret, U., Zehe, E., Wulfmeyer, V., Warrach-Sagi, K., and Liebert, J.: HESS Opinions "Should we apply bias correction to global and regional climate model data?", Hydrology and Earth System Sciences, 16, 3391–3404, doi:10.5194/hess-16-3391-2012, 2012.

Fatichi, S., Rimkus, S., Burlando, P., and Bordoy, R.: Does internal climate variability overwhelm climate change signals in streamflow? The upper Po and Rhone basin case studies., The Science of the Total Environment, doi:10.1016/j.scitotenv.2013.12.014, 2014.

Fenicia, F., Kavetski, D., Savenije, H. H. G., Clark, M. P., Schoups, G., Pfister, L., and Freer, J.: Catchment properties, function, and conceptual model representation: is there a correspondence?, Hydrological Processes, 28, 2451–2467, doi:10.1002/hyp.9726, 2014.

Finger, D., Heinrich, G., Gobiet, A., and Bauder, A.: Projections of future water resources and their uncertainty in a glacierized catchment in the Swiss Alps and the subsequent effects on hydropower production during the 21st century, Water Resources Research, 48, doi:10.1029/2011WR010733, 2012.

Fischer, A. M., Weigel, A. P., Buser, C. M., Knutti, R., Künsch, H. R., Liniger, M. A., Schär, C., and Appenzeller, C.: Climate change projections for Switzerland based on a Bayesian multi-model approach, International Journal of Climatology, 32, 2348–2371, doi:10.1002/joc.3396, 2012.

Fischer, E. M. and Knutti, R.: Robust projections of combined humidity and temperature extremes, Nature Climate Change, 3, 126–130, doi:10.1038/nclimate1682, 2012.

Fischer, E. M. and Knutti, R.: Detection of spatially aggregated changes in temperature and precipitation extremes, Geophysical Research Letters, 41, doi:10.1002/2013GL058499, 2014.

Fischer, E. M. and Schär, C.: Future changes in daily summer temperature variability: driving processes and role for temperature extremes, Climate Dynamics, 33, 917–935, doi:10.1007/s00382-008-0473-8, 2009.

Fischer, E. M., Seneviratne, S. I., Vidale, P. L., Lüthi, D., and Schär, C.: Soil Moisture-Atmosphere Interactions during the 2003 European Summer Heat Wave, Journal of Climate, 20, 5081–5099, doi:10.1175/JCLI4288.1, 2007.

Fischer, E. M., Beyerle, U., and Knutti, R.: Robust spatially aggregated projections of climate extremes, Nature Climate Change, 3, 1033–1038, doi:10.1038/nclimate2051, 2013.

Fischer, M., Huss, M., Barboux, C., and Hoelzle, M.: The new Swiss Glacier Inventory SGI2010: Relevance of using high-resolution source data in areas dominated by very small glaciers, Arctic, Antarctic and Alpine Research, 46, 933–945, doi:10.1657/1938-4246-46.4.933, 2014.

Fisher, P.: Models of uncertainty in spatial data, in: Geographical Information Systems: Principles and Technical Issues, edited by Longley, P. A., Goodchild, M. F., and Maguire, D. J., pp. 191–205, John Wiley, New York, 2nd ed edn., 1999.

Forzieri, G., Feyen, L., Rojas, R., Flörke, M., Wimmer, F., and Bianchi, A.: Ensemble projections of future streamflow droughts in Europe, Hydrology and Earth System Sciences, 18, 85–108, doi:10.5194/hess-18-85-2014, 2014.

Fowler, H. J., Blenkinsop, S., and Tebaldi, C.: Linking climate change modelling to impacts studies : recent advances in downscaling techniques for hydrological, International Journal of Climatology, 27, 1547–1578, doi:10.1002/joc, 2007.

Frei, C.: Interpolation of temperature in a mountainous region using nonlinear profiles and non-Euclidean distances, International Journal of Climatology, 34, 1585–1605, doi:10.1002/joc.3786, 2013.

Frei, C. and Schär, C.: A precipitation climatology of the Alps from high-resolution rain-gauge observations, International Journal of Climatology, 18, 873–900, doi:10.1002/(SICI)1097-0088(19980630)18:8<873::AID-JOC255>3.0.CO;2-9, 1998.

Frei, C., Christensen, J., Déqué, M., Jacob, D., Jones, R., and Vidale, P. L.: Daily precipitation statistics in regional climate models: Evaluation and intercomparison for the European Alps, Journal of Geophysical Research, 108, 1–19, doi:10.1029/2002JD002287, 2003.

Frei, C., Schöll, R., Fukutome, S., Schmidli, J., and Vidale, P. L.: Future change of precipitation extremes in Europe: Intercomparison of scenarios from regional climate models, Journal of Geophysical Research, 111, D06 105, doi:10.1029/2005JD005965, 2006.

Füssel, H.-M. and Klein, R. J. T.: Climate Change Vulnerability Assessments: An Evolution of Conceptual Thinking, Climatic Change, 75, 301–329, doi:10.1007/s10584-006-0329-3, 2006.

Gebrehiwot, S. G., Seibert, J., Gärdenäs, A. I., Mellander, P.-E., and Bishop, K.: Hydrological change detection using modeling: Half a century of runoff from four rivers in the Blue Nile Basin, Water Resources Research, 49, 3842–3851, doi:10.1002/wrcr.20319, 2013.

Gharari, S., Hrachowitz, M., Fenicia, F., and Savenije, H. H. G.: An approach to identify time consistent model parameters: sub-period calibration, Hydrology and Earth System Sciences, 17, 149–161, doi:10.5194/hess-17-149-2013, 2013.

Gómez-Navarro, J. J., Montávez, J. P., Jerez, S., Jiménez-Guerrero, P., and Zorita, E.: What is the role of the observational dataset in the evaluation and scoring of climate models?, Geophysical Research Letters, 39, doi:10.1029/2012GL054206, 2012.

Gornish, E. S., Hamilton, J. A., Barberán, A., Benito, B. M., Binzer, A., DeMeester, J. E., Gruwez, R., Moreira, B., Taheri, S., Tomiolo, S., Vinagre, C., Vuarin, P., and Weaver, J.: Interdisciplinary Climate Change Collaborations Are Essential for Early-Career Scientists, Eos, Transactions American Geophysical Union, 94, 151–151, doi:10.1002/2013EO160003, 2013.

Greve, P., Orlowsky, B., Mueller, B., Sheffield, J., Reichstein, M., and Seneviratne, S. I.: Global assessment of trends in wetting and drying over land, Nature Geoscience, 7, 716–721, doi:10.1038/ngeo2247, 2014.

Gudmundsson, L., Bremnes, J. B., Haugen, J. E., and Engen-Skaugen, T.: Technical Note: Downscaling RCM precipitation to the station scale using statistical transformations - a comparison of methods, Hydrology and Earth System Sciences, 16, 3383–3390, doi:10.5194/hess-16-3383-2012, 2012.

Gupta, H. V., Clark, M. P., Vrugt, J. A., Abramowitz, G., and Ye, M.: Towards a comprehensive assessment of model structural adequacy, Water Resources Research, 48, doi:10.1029/2011WR011044, 2012.

Gupta, H. V., Perrin, C., Blöschl, G., Montanari, A., Kumar, R., Clark, M., and Andréassian, V.: Large-sample hydrology: a need to balance depth with breadth, Hydrology and Earth System Sciences, 18, 463–477, doi:10.5194/hess-18-463-2014, 2014.

Hagemann, S., Chen, C., Haerter, J. O., Heinke, J., Gerten, D., and Piani, C.: Impact of a Statistical Bias Correction on the Projected Hydrological Changes Obtained from Three GCMs and Two Hydrology Models, Journal of Hydrometeorology, 12, 556–578, doi:10.1175/2011JHM1336.1, 2011.

Harlin, J. and Kung, C.-S.: Parameter uncertainty and simulation of design floods in Sweden, Journal of Hydrology, 137, 209–230, doi:10.1016/0022-1694(92)90057-3, 1992.

Hartmann, A., Barberá, J. A., Lange, J., Andreo, B., and Weiler, M.: Progress in the hydrologic simulation of time variant recharge areas of karst systems - Exemplified at a karst spring in Southern Spain, Advances in Water Resources, 54, 149–160, doi:10.1016/j.advwatres.2013.01.010, 2013.

Hawkins, E. and Sutton, R.: The Potential to Narrow Uncertainty in Regional Climate Predictions, Bulletin of the American Meteorological Society, 90, 1095–1107, doi:10.1175/2009BAMS2607.1, 2009.

Hawkins, E. and Sutton, R.: The potential to narrow uncertainty in projections of regional precipitation change, Climate Dynamics, 37, 407–418, doi:10.1007/s00382-010-0810-6, 2010.

Hawkins, E. and Sutton, R.: Time of emergence of climate signals, Geophysical Research Letters, 39, 1–6, doi:10.1029/2011GL050087, 2012.

Haylock, M. R., Hofstra, N., Klein Tank, a. M. G., Klok, E. J., Jones, P. D., and New, M.: A European daily high-resolution gridded data set of surface temperature and precipitation for 1950-2006, Journal of Geophysical Research, 113, D20 119, doi:10.1029/2008JD010201, 2008.

Hazell, P., Anderson, J., Balzer, N., Hastrup Clemmensen, A., Hess, U., and Rispoli, F.: The Potential for Scale and Sustainability The Potential for Scale and Sustainability in Weather Index Insurance, Tech. rep., Rome: International Fund for Agricultural Development and World Food Programme, 2010.

Herman, J. D., Reed, P. M., and Wagener, T.: Time-varying sensitivity analysis clarifies the effects of watershed model formulation on model behavior, Water Resources Research, 49, 1400–1414, doi:10.1002/wrcr.20124, 2013.

Hock, R.: A distributed temperature-index ice- and snowmelt model including potential direct solar radiation, Journal of Glaciology, 45, 101–111, 1999.

Hofstra, N., Haylock, M., New, M., and Jones, P. D.: Testing E-OBS European high-resolution gridded data set of daily precipitation and surface temperature, Journal of Geophysical Research, 114, D21 101, doi:10.1029/2009JD011799, 2009.

Hoppe, R., Wesselink, A., and Cairns, R.: Lost in the problem: The role of boundary organisations in the governance of climate change, Wiley Interdisciplinary Reviews: Climate Change, 4, 283–300, doi:10.1002/wcc.225, 2013.

Horton, P., Schaefli, B., Mezghani, A., Hingray, B., and Musy, A.: Assessment of climate-change impacts on alpine discharge regimes with climate model uncertainty, Hydrological Processes, 20, 2091–2109, doi:10.1002/hyp.6197, 2006.

Hrachowitz, M., Savenije, H., Blöschl, G., McDonnell, J., Sivapalan, M., Pomeroy, J., Arheimer, B., Blume, T., Clark, M. P., Ehret, U., Fenicia, F., Freer, J. E., Gelfan, A., Gupta, H., Hughes, D., Hut, R., Montanari, A., Pande, S., Tetzlaff, D., Troch, P. a., Uhlenbrook, S., Wagener, T., Winsemius, H., Woods, R., Zehe, E., and Cudennec, C.: A decade of Predictions in Ungauged Basins (PUB) - a review, Hydrological Sciences Journal, 58, 1198–1255, doi:10.1080/02626667.2013.803183, 2013.

Huggel, C., Khabarov, N., Obersteiner, M., and Ramírez, J. M.: Implementation and integrated numerical modeling of a landslide early warning system: a pilot study in Colombia, Natural Hazards, 52, 501–518, doi:10.1007/s11069-009-9393-0, 2010.

Huss, M. and Farinotti, D.: Distributed ice thickness and volume of all glaciers around the globe, Journal of Geophysical Research, 117, F04 010, doi:10.1029/2012JF002523, 2012.

Huss, M., Hock, R., Bauder, A., and Funk, M.: 100-year mass changes in the Swiss Alps linked to the Atlantic Multidecadal Oscillation, Geophysical Research Letters, 37, doi:10.1029/2010GL042616, 2010a.

Huss, M., Jouvet, G., Farinotti, D., and Bauder, A.: Future high-mountain hydrology: a new parameterization of glacier retreat, Hydrology and Earth System Sciences, 14, 815–829, doi:10.5194/hess-14-815-2010, 2010b.

Huss, M., Zemp, M., Joerg, P. C., and Salzmann, N.: High uncertainty in 21st century runoff projections from glacierized basins, Journal of Hydrology, 510, 35–48, doi:10.1016/j.jhydrol.2013.12.017, 2013.

Huth, R., Beck, C., Philipp, A., Demuzere, M., Ustrnul, Z., Cahynová, M., Kyselý, J., and Tveito, O. E.: Classifications of atmospheric circulation patterns: Recent advances and applications, Annals of the New York Academy of Sciences, 1146, 105–152, doi:10.1196/annals.1446.019, 2008.

IISD: CRiSTAL User'Äôs Manual Version 5 Community-based Risk Screening Tool - Adaptation and Livelihoods, Tech. rep., The International Institute for Sustainable Development, 2012.

Immerzeel, W., Petersen, L., Ragettli, S., and Pellicciotti, F.: The importance of observed gradients of air temperature and precipitation for modeling runoff froma glacierized watershed in the Nepalese Himalayas, Water Resources Research, 50, 2212–2226, doi:10.1002/2013WR014506, 2014.

IPCC: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge, United Kingdom and New York, NY, USA, doi:10.1017/CBO9781107415324, 2013.

Isotta, F. A., Frei, C., Weilguni, V., Perčec Tadić, M., Lassègues, P., Rudolf, B., Pavan, V., Cacciamani, C., Antolini, G., Ratto, S. M., Munari, M., Micheletti, S., Bonati, V., Lussana, C., Ronchi, C., Panettieri, E., Marigo, G., and Vertačnik, G.: The climate of daily precipitation in the Alps: development and analysis of a high-resolution grid dataset from pan-Alpine rain-gauge data, International Journal of Climatology, 34, doi:10.1002/joc.3794, 2014.

Jakeman, A. and Hornberger, G.: How Much Complexity Is Warranted in a Rainfall-Runoff Model ?, Water Resources Research, 29, 2637–2649, doi:10.1029/93WR00877, 1993.

James, P. M.: An assessment of European synoptic variability in Hadley Centre Global Environmental models based on an objective classification of weather regimes, Climate Dynamics, 27, 215–231, doi:10.1007/s00382-006-0133-9, 2006.

Jaun, S., Ahrens, B., Walser, A., Ewen, T., and Schär, C.: A probabilistic view on the August 2005 floods in the upper Rhine catchment, Nat. Hazard Earth Sys., 8, 281–291, doi:10.5194/nhess-8-281-2008, 2008.

Jiménez Cisneros, B., Oki, T., Arnell, N., Benito, G., Cogley, J. G., Döll, P., Jiang, T., and Mwakalila, S. S.: Freshwater Resources, in: Climate Change 2014: Impacts, Adaptation, and Vulnerability. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, October 2013, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2014.

Jung, M., Reichstein, M., Ciais, P., Seneviratne, S. I., Sheffield, J., Goulden, M. L., Bonan, G., Cescatti, A., Chen, J., de Jeu, R., Dolman, a. J., Eugster, W., Gerten, D., Gianelle, D., Gobron, N., Heinke, J., Kimball, J., Law, B. E., Montagnani, L., Mu, Q., Mueller, B., Oleson, K., Papale, D., Richardson, A. D., Roupsard, O., Running, S., Tomelleri, E., Viovy, N., Weber, U., Williams, C., Wood, E., Zaehle, S., and Zhang, K.: Recent decline in the global land evapotranspiration trend due to limited moisture supply., Nature, 467, 951–4, doi:10.1038/nature09396, 2010.

Kelly, P. and Adger, W.: Theory and practice in assessing vulnerability to climate change and facilitating adaptation, Climatic Change, 47, 325–352, 2000.

Kerkhoff, C., Künsch, H. R., and Schär, C.: Assessment of Bias Assumptions for Climate Models, Journal of Climate, 27, 6799–6818, doi:10.1175/JCLI-D-13-00716.1, 2014.

Kirchner, J. W.: Getting the right answers for the right reasons: Linking measurements, analyses, and models to advance the science of hydrology, Water Resources Research, 42, doi:10.1029/2005WR004362, 2006.

Kleiner, J.: Uncertainty communication - The use of visualisations to provide uncertainty information, Ph.D. thesis, Master Thesis, University of Zurich, 2013.

Klemeš, V.: Operational testing of hydrological simulation models, Hydrological Sciences Journal, 31, 13–24, doi:10.1080/02626668609491024, 1986.

Knutti, R. and Sedláček, J.: Robustness and uncertainties in the new CMIP5 climate model projections, Nature Climate Change, 3, 369–373, doi:10.1038/nclimate1716, 2012.

Knutti, R., Furrer, R., Tebaldi, C., Cermak, J., and Meehl, G. A.: Challenges in Combining Projections from Multiple Climate Models, Journal of Climate, 23, 2739–2758, doi:10.1175/2009JCLI3361.1, 2010.

Knutti, R., Masson, D., and Gettelman, A.: Climate model genealogy: Generation CMIP5 and how we got there, Geophysical Research Letters, 40, 1194–1199, doi:10.1002/grl.50256, 2013.

Konz, M. and Seibert, J.: On the value of glacier mass balances for hydrological model calibration, Journal of Hydrology, 385, 238–246, doi:10.1016/j.jhydrol.2010.02.025, 2010.

Köplin, N., Schädler, B., Viviroli, D., and Weingartner, R.: Relating climate change signals and physiographic catchment properties to clustered hydrological response types, Hydrology and Earth System Sciences, 16, 2267–2283, doi:10.5194/hess-16-2267-2012, 2012.

Köplin, N., Schädler, B., Viviroli, D., and Weingartner, R.: The importance of glacier and forest change in hydrological climate-impact studies, Hydrology and Earth System Sciences, 17, 619–635, doi:10.5194/hess-17-619-2013, 2013.

Kotlarski, S., Keuler, K., Christensen, O. B., Colette, A., Déqué, M., Gobiet, A., Goergen, K., Jacob, D., Lüthi, D., van Meijgaard, E., Nikulin, G., Schär, C., Teichmann, C., Vautard, R., Warrach-Sagi, K., and Wulfmeyer, V.: Regional climate modeling on

European scales: a joint standard evaluation of the EURO-CORDEX RCM ensemble, Geoscientific Model Development, 7, 1297–1333, doi:10.5194/gmd-7-1297-2014, 2014.

Kyselý, J. and Plavcová, E.: Biases in the diurnal temperature range in Central Europe in an ensemble of regional climate models and their possible causes, Climate Dynamics, doi:10.1007/s00382-011-1200-4, 2011.

Lang, H.: Forecasting Meltwater Runoff from Snow-Covered Areas and from Glacier Basins, River flow modelling and forecasting, 3, 99–127, doi:10.1007/978-94-009-4536-4_5, 1986.

Lemos, M. C. and Morehouse, B. J.: The co-production of science and policy in integrated climate assessments, Global Environmental Change, 15, 57–68, doi:10.1016/j.gloenvcha.2004.09.004, 2005.

Lemos, M. C. and Rood, R. B.: Climate projections and their impact on policy and practice, Wiley Interdisciplinary Reviews: Climate Change, 1, 670–682, doi:10.1002/wcc.71, 2010.

Lemos, M. C., Kirchhoff, C. J., Kalafatis, S. E., Scavia, D., and Rood, R. B.: Moving Climate Information off the Shelf: Boundary Chains and the Role of RISAs as Adaptive Organizations, Weather, Climate, and Society, 6, 273–285, doi:10.1175/WCAS-D-13-00044.1, 2014.

Li, C., Sinha, E., Horton, D. E., Diffenbaugh, N. S., and Michalak, A. M.: Joint bias correction of temperature and precipitation in climate model simulations, Journal of Geophysical Research - Atmospheres, 119, 153–162, doi:10.1002/ 2014JD022514, 2014.

Liang, X., Lettenmaier, D. P., Wood, E. F., and Burges, S. J.: A simple hydrologically based model of land surface water and energy fluxes for general circulation models, Journal of Geophysical Research, 99, 14 415, doi:10.1029/94JD00483, 1994.

Lindström, G., Johansson, B., Persson, M., Gardelin, M., and Bergström, S.: Development and test of the distributed HBV-96 hydrological model, Journal of Hydrology, 201, 272–288, doi:10.1016/S0022-1694(97)00041-3, 1997.

Locke, S.: Golem science and the public understanding of science: from deficit to dilemma, Public Understanding of Science, 8, 75–92, 1999.

Lombardot, T., Kottmann, R., Giuliani, G., de Bono, A., Addor, N., and Glöckner, F. O.: MetaLook: a 3D visualisation software for marine ecological genomics, BMC Bioinformatics, 8, doi:10.1186/1471-2105-8-406, 2007.

MacEachren, A.: Visualizing uncertain information, Cartographic Perspectives, pp. 10–19, 1992.

MacEachren, A. M., Robinson, A., Hopper, S., Gardner, S., Murray, R., Gahegan, M., and Hetzler, E.: Visualizing Geospatial Information Uncertainty: What We Know and What We Need to Know, Cartography and Geographic Information Science, 32, 139–160, doi:10.1559/1523040054738936, 2005.

Mahlstein, I., Portmann, R. W., Daniel, J. S., Solomon, S., and Knutti, R.: Perceptible changes in regional precipitation in a future climate, Geophysical Research Letters, 39, doi:10.1029/2011GL050738, 2012.

Maraun, D.: Nonstationarities of regional climate model biases in European seasonal mean temperature and precipitation sums, Geophysical Research Letters, 39, L06 706, doi:10.1029/2012GL051210, 2012.

Maraun, D., Wetterhall, F., Ireson, A. M., Chandler, R. E., Kendon, E. J., Widmann, M., Brienen, S., Rust, H. W., Sauter, T., Themel, M., Venema, V. K. C., Chun, K. P., Goodess, C. M., Jones, R. G., Onof, C., Vrac, M., and Thiele-Eich, I.: Precipitation downscaling under climate change: Recent developments to bridge the gap between dynamical models and the end user, Reviews of Geophysics, 48, 1–34, doi:10.1029/2009RG000314, 2010.

Maraun, D., Osborn, T. J., and Rust, H. W.: The influence of synoptic airflow on UK daily precipitation extremes. Part II: regional climate model and E-OBS data validation, Climate Dynamics, doi:10.1007/s00382-011-1176-0, 2011.

Masson, D. and Knutti, R.: Climate model genealogy, Geophysical Research Letters, 38, L08 703, doi:10.1029/2011GL046864, 2011.

Mastrandrea, M. D., Field, C. B., Stocker, T. F., Edenhofer, O., Ebi, K. L., Frame, D. J., Held, H., Kriegler, E., Mach, K. J., Matschoss, P. R., Plattner, G.-K., Yohe, G. W., and Zwiers, F. W.: Guidance Note for Lead Authors of the IPCC Fifth Assessment Report on Consistent Treatment of Uncertainties, Intergovernmental Panel on Climate Change (IPCC), URL http://193.194.138.236/pdf/supporting-material/uncertainty-guidance-note.pdf, 2010.

McNie, E. C.: Reconciling the supply of scientific information with user demands: an analysis of the problem and review of the literature, Environmental Science & Policy, 10, 17–38, doi:10.1016/j.envsci.2006.10.004, 2007.

Mearns, L. O., Gutowski, W. J., Jones, R., Leung, L.-Y., McGinnis, S., Nunes, A. M. B., and Qian:, Y.: A regional climate change assessment program, EOS, 90, 311–312, doi:10.1029/2009EO360002, 2009.

Meinshausen, M., Smith, S. J., Calvin, K., Daniel, J. S., Kainuma, M. L. T., Lamarque, J.-F., Matsumoto, K., Montzka, S. A., Raper, S. C. B., Riahi, K., Thomson, A., Velders, G. J. M., and Vuuren, D. P.: The RCP greenhouse gas concentrations and their extensions from 1765 to 2300, Climatic Change, 109, 213–241, doi:10.1007/s10584-011-0156-z, 2011.

Merta, M., Seidler, C., Bianchin, S., Heilmeier, H., and Richert, E.: Analysis of Land Use Change in the Eastern Ore Mts. Regarding Both Nature Protection and Flood Prevention, Soil and Water Res., 3, 105–115, 2008.

Merz, R., Parajka, J., and Blöschl, G.: Time stability of catchment model parameters: Implications for climate impact analyses, Water Resources Research, 47, doi:10.1029/2010WR009505, 2011.

MeteoSwiss: Documentation of MeteoSwiss Grid-Data Products - Monthly and Yearly Precipitation : RhiresM and RhiresY, Tech. rep., Federal Office of Meteorology and Climatology, MeteoSwiss, URL http://www.meteoschweiz.admin.ch/web/de/services/datenportal/gitterdaten/precip.Par.0006.DownloadFile.tmp/precipitationproducts.pdf, 2013a.

MeteoSwiss: Documentation of MeteoSwiss Grid-Data Products - Monthly and Yearly Mean Temperature: TabsM and TabsY, Tech. rep., Federal Office of Meteorology and Climatology, MeteoSwiss, URL http://www.meteoschweiz.admin.ch/web/de/services/datenportal/gitterdaten/temp/tabsm.Par.0004.DownloadFile.tmp/proddoctabsm.pdf, 2013b.

Morice, C. P., Kennedy, J. J., Rayner, N. a., and Jones, P. D.: Quantifying uncertainties in global and regional temperature change using an ensemble of observational estimates: The HadCRUT4 data set, Journal of Geophysical Research, 117, D08 101, doi:10.1029/2011JD017187, 2012.

Moss, R. H., Edmonds, J. a., Hibbard, K. a., Manning, M. R., Rose, S. K., van Vuuren, D. P., Carter, T. R., Emori, S., Kainuma, M., Kram, T., Meehl, G. a., Mitchell, J. F. B., Nakicenovic, N., Riahi, K., Smith, S. J., Stouffer, R. J., Thomson, A. M., Weyant, J. P., and Wilbanks, T. J.: The next generation of scenarios for climate change research and assessment., Nature, 463, 747–56, doi:10.1038/nature08823, 2010.

Mueller, B. and Seneviratne, S.: Systematic land climate and evapotranspiration biases in CMIP5 simulations, Geophysical Research Letters, 41, 128–134, doi:10.1002/2013GL058055, 2014.

Muerth, M. J., Gauvin St-Denis, B., Ricard, S., Velázquez, J. A., Schmid, J., Minville, M., Caya, D., Chaumont, D., Ludwig, R., and Turcotte, R.: On the need for bias correction in regional climate scenarios to assess climate change impacts on river runoff, Hydrology and Earth System Sciences, 17, 1189–1204, doi:10.5194/hess-17-1189-2013, 2013.

Nakicenovic, N. and Swart, R.: IPCC Special Report Emission Scenarios, Cambridge University Press, UK, 2000.

Nash, J. and Sutcliffe, J.: River flow forecasting through conceptual models. Part I: a discussion of principles, Journal of Hydrology, 10, 282–290, 1970.

Neukum, C. and Azzam, R.: Impact of climate change on groundwater recharge in a small catchment in the Black Forest, Germany, Hydrogeology Journal, 20, 547–560, doi:10.1007/s10040-011-0827-x, 2012.

Noguer, M., Jones, R., and Murphy, J.: Sources of systematic errors in the climatology of a regional climate model over Europe, Climate Dynamics, 14, 691–712, doi:10.1007/s003820050249, 1998.

Nowotny, H., Scott, P., and Gibbons, M. T.: Re-Thinking Science: Knowledge and the Public in an Age of Uncertainty, Polity Press, Cambridge, 2001.

O'Brien, K., Eriksen, S., Nygaard, L. P., and Schjolden, A.: Why different interpretations of vulnerability matter in climate change discourses, Climate Policy, 7, 73–88, doi:10.1080/14693062.2007.9685639, 2007.

Oudin, L., Hervieu, F., Michel, C., Perrin, C., Andréassian, V., Anctil, F., and Loumagne, C.: Which potential evapotranspiration input for a lumped rainfall-runoff model?, Journal of Hydrology, 303, 290–306, doi:10.1016/j.jhydrol.2004.08.026, 2005.

Pappenberger, F. and Beven, K. J.: Ignorance is bliss: Or seven reasons not to use uncertainty analysis, Water Resources Research, 42, 1–8, doi:10.1029/2005WR004820, 2006.

Parajka, J. and Blöschl, G.: The value of MODIS snow cover data in validating and calibrating conceptual hydrologic models, Journal of Hydrology, 358, 240–258, doi:10.1016/j.jhydrol.2008.06.006, 2008.

Perez, J., Menendez, M., Mendez, F. J., and Losada, I. J.: Evaluating the performance of CMIP3 and CMIP5 global climate models over the north-east Atlantic region, Climate Dynamics, 43, 1–18, doi:10.1007/s00382-014-2078-8, 2014.

Perica, S. and Foufoula-Georgiou, E.: Model for multiscale disaggregation of spatial rainfall based on coupling meteorological and scaling descriptions, Journal of Geophysical Research, 101, 26 347–26 361, doi:10.1029/96JD01870, 1996.

Perkins, S. E. and Fischer, E. M.: The usefulness of different realizations for the model evaluation of regional trends in heat waves, Geophysical Research Letters, 40, 5793–5797, doi:10.1002/2013GL057833, 2013.

Perrin, C. and Michel, C.: Does a large number of parameters enhance model performance? Comparative assessment of common catchment model structures on 429 catchments, Journal of Hydrology, 242, doi:10.1016/S0022-1694(00)00393-0, 2001.

Philipp, A., Bartholy, J., Beck, C., Erpicum, M., Esteban, P., Fettweis, X., Huth, R., James, P., Jourdain, S., and Kreienkamp, F.: Cost733cat - A database of weather and circulation type classifications, Physics and Chemistry of the Earth, Parts A/B/C, 35, 360–373, doi:10.1016/j.pce.2009.12.010, 2010.

Philipp, A., Beck, C., Huth, R., and Jacobeit, J.: Development and comparison of circulation type classifications using the COST 733 dataset and software, International Journal of Climatology, doi:10.1002/joc.3920, 2014.

Piani, C., Weedon, G., Best, M., Gomes, S., Viterbo, P., Hagemann, S., and Haerter, J.: Statistical bias correction of global simulated daily precipitation and temperature for the application of hydrological models, Journal of Hydrology, 395, 199–215, doi:10.1016/j.jhydrol.2010.10.024, 2010.

Plavcová, E. and Kyselý, J.: Evaluation of daily temperatures in Central Europe and their links to large-scale circulation in an ensemble of regional climate models, Tellus A, 63, 763–781, doi:10.1111/j.1600-0870.2011.00514.x, 2011.

Poulin, A., Brissette, F., Leconte, R., Arsenault, R., and Malo, J.-S.: Uncertainty of hydrological modelling in climate change impact studies in a Canadian, snow-dominated river basin, Journal of Hydrology, 409, 626–636, doi:10.1016/j.jhydrol.2011.08.057, 2011.

Prudhomme, C. and Davies, H.: Assessing uncertainties in climate change impact analyses on the river flow regimes in the UK. Part 1: baseline climate, Climatic Change, 93, 177–195, doi:10.1007/s10584-008-9464-3, 2008.

Prudhomme, C. and Davies, H.: Assessing uncertainties in climate change impact analyses on the river flow regimes in the UK. Part 2: future climate, Climatic Change, 93, 197–222, doi:10.1007/s10584-008-9461-6, 2009.

Prudhomme, C. and Genevier, M.: Can atmospheric circulation be linked to flooding in Europe?, Hydrological Processes, 25, 1180–1190, doi:10.1002/hyp.7879, 2011.

Prudhomme, C., Wilby, R., Crooks, S., Kay, A., and Reynard, N.: Scenario-neutral approach to climate change impact studies: Application to flood risk, Journal of Hydrology, 390, 198–209, doi:10.1016/j.jhydrol.2010.06.043, 2010.

R Core Team: R: A Language and Environment for Statistical Computing, URL http://www.r-project.org/, 2014.

Racherla, P. N., Shindell, D. T., and Faluvegi, G. S.: The added value to global model projections of climate change by dynamical downscaling: A case study over the continental U.S. using the GISS-ModelE2 and WRF models, Journal of Geophysical Research: Atmospheres, 117, doi:10.1029/2012JD018091, 2012.

Rauscher, S. a., Coppola, E., Piani, C., and Giorgi, F.: Resolution effects on regional climate model simulations of seasonal precipitation over Europe, Climate Dynamics, 35, 685–711, doi:10.1007/s00382-009-0607-7, 2010.

Refsgaard, J. C., Madsen, H., Andréassian, V., Arnbjerg-Nielsen, K., Davidson, T. a., Drews, M., Hamilton, D. P., Jeppesen, E., Kjellström, E., Olesen, J. E., Sonnenborg, T. O., Trolle, D., Willems, P., and Christensen, J. H.: A framework for testing the ability of models to project climate change and its impacts, Climatic Change, 122, 271–282, doi:10.1007/s10584-013-0990-2, 2013.

Riediger, U. and Gratzki, A.: Future weather types and their influence on mean and extreme climate indices for precipitation and temperature in Central Europe, Meteorologische Zeitschrift, 23, 231–252, doi:10.1127/0941-2948/2014/0519, 2014.

Rohrer, M.: Climate Change and Circulation Types in the Alpine Region, Tech. rep., Scientific Report MeteoSwiss 91, Scientific Report MeteoSwiss 91, URL http://www.meteoschweiz.admin.ch/content/dam/meteoswiss/en/Ungebundene-Seiten/Publikationen/Scientific-Reports/doc/sr91rohrer.pdf, 2013.

Rojas, R., Feyen, L., Dosio, A., and Bavera, D.: Improving pan-European hydrological simulation of extreme events through statistical bias correction of RCM-driven climate simulations, Hydrology and Earth System Sciences, 15, 2599–2620, doi:10.5194/hess-15-2599-2011, 2011.

Rössler, O., Diekkrüger, B., and Löffler, J.: Potential drought stress in a Swiss mountain catchment-Ensemble forecasting of high mountain soil moisture reveals a drastic decrease, despite major uncertainties, Water Resources Research, 48, W04 521, doi:10.1029/2011WR011188, 2012.

Rössler, O., Addor, N., Bernhard, L., Figura, S., Köplin, N., Livingstone, D., Schädler, B., Seibert, J., and Weingartner, R.: Hydrological responses to climate change : river runoff and groundwater, in: Toward quantitative scenarios of climate change impacts in Switzerland, edited by CH2014-Impacts, chap. Chapter 6, FOEN, MeteoSwiss, C2SM, Agroscope, and ProClim, Bern, Switzerland, 2014.

Saldaña, J.: The Coding Manual for Qualitative Researchers, Sage Publications, London, 2nd ed edn., 2013.

Salzmann, N., Kotlarski, S., von Waldow, H., and Rajczak, J.: Advancing and facilitating the use of RCM data in climate impacts research, in: Impacts World 2013 Conference Proceedings, pp. 565–572, Potsdam: Potsdam Institute for Climate Impact Research, doi:10.2312/pik.2013.001, 2013.

Savenije, H. H. G.: "The art of hydrology", Hydrology and Earth System Sciences, 13, 157–161, doi:10.5194/hess-13-157-2009, 2009.

Schaefli, B. and Gupta, H. V.: Do Nash values have value?, Hydrological Processes, 21, 2075–2080, doi:10.1002/hyp.6825, 2007.

Schär, C., Vidale, P., Lüthi, D., and Frei, C.: The role of increasing temperature variability in European summer heatwaves, Nature, 427, 3926–3928, doi:10.1038/nature02230.1., 2004.

Schiemann, R. and Frei, C.: How to quantify the resolution of surface climate by circulation types: An example for Alpine precipitation, Physics and Chemistry of the Earth, Parts A/B/C, 35, 403–410, doi:10.1016/j.pce.2009.09.005, 2010.

Schmocker-Fackel, P. and Naef, F.: More frequent flooding? Changes in flood frequency in Switzerland since 1850, Journal of Hydrology, 381, 1–8, doi:10.1016/j.jhydrol.2009.09.022, 2010.

Schulla, J.: Model Description WaSiM, Hydrology Software Consulting J. Schulla, Switzerland, URL <http://www.wasim.ch>, 2013.

Schwarb, M.: The Alpine precipitation climate: Evaluation of a high-resolution analysis scheme using comprehensive rain-gauge data, Phd thesis 13911, ETH, PhD thesis, 2000.

SDC Swiss Agency for Development and Cooperation: CEDRIG: Climate, Environment and Disaster Risk Reduction Integration Guidance, Part II CEDRIG Handbook, Tech. rep., Berne, Switzerland, URL http://www.sdc-drr.net/system/files/CEDRIG_Part_II_Handbook_EN_Web.pdf, 2012.

Seibert, J.: Multi-criteria calibration of a conceptual runoff model using a genetic algorithm, Hydrology and Earth System Sciences, 4, 215–224, doi:10.5194/hess-4-215-2000, 2000.

Seibert, J.: Reliability of Model Predictions Outside Calibration Conditions, Nordic Hydrology, 34, 477–492, doi:10.2166/nh.2003.028, 2003.

Seibert, J. and McDonnell, J. J.: On the dialog between experimentalist and modeler in catchment hydrology: Use of soft data for multicriteria model calibration, Water Resources Research, 38, 1241, doi:10.1029/2001WR000978, 2002.

Seibert, J. and Vis, M. J. P.: Teaching hydrological modeling with a user-friendly catchment-runoff-model software package, Hydrology and Earth System Sciences, 16, 3315–3325, doi:10.5194/hess-16-3315-2012, 2012.

Seiller, G., Anctil, F., and Perrin, C.: Multimodel evaluation of twenty lumped hydrological models under contrasted climate conditions, Hydrology and Earth System Sciences, 16, 1171–1189, doi:10.5194/hess-16-1171-2012, 2012.

Seneviratne, S. I., Lehner, I., Gurtz, J., Teuling, A. J., Lang, H., Moser, U., Grebner, D., Menzel, L., Schroff, K., Vitvar, T., and Zappa, M.: Swiss prealpine Rietholzbach research catchment and lysimeter: 32 year time series and 2003 drought event, Water Resources Research, 48, W06 526, doi:10.1029/2011WR011749, 2012a.

Seneviratne, S. I., Nicholls, N., Easterling, D., Goodess, C. M., Kanae, S., Kossin, J., Luo, Y., Marengo, J., McInnes, K., Rahimi, M., Reichstein, M., Sorteberg, A., Vera, C., and Zhang, X.: Changes in climate extremes and their impacts on the natural physical environment, in: Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation, edited by Field, C. B., Barros, V., Stocker, T. F., Qin, D., Dokken, D. J., Ebi, K. L., Mastrandrea, M. D., Mach, K. J., Plattner, G. K., Allen, S. K., Tignor, M., and Midgley, P. M., pp. 109–230, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2012b.

Sheffield, J., Wood, E. F., and Roderick, M. L.: Little change in global drought over the past 60 years., Nature, 491, 435–8, doi:10.1038/nature11575, 2012.

Shepherd, T. G.: Atmospheric circulation as a source of uncertainty in climate change projections, Nature Geoscience, 7, 703–708, doi:10.1038/ngeo2253, 2014.

Singh, N., Abiven, S., Torn, M. S., and Schmidt, M. W. I.: Fire-derived organic carbon in soil turns over on a centennial scale, Biogeosciences, 9, 2847–2857, doi:10.5194/bg-9-2847-2012, 2012.

Singh, R., Wagener, T., van Werkhoven, K., Mann, M. E., and Crane, R.: A trading-space-for-time approach to probabilistic continuous streamflow predictions in a changing climate - accounting for changing watershed behavior, Hydrology and Earth System Sciences, 15, 3591–3603, doi:10.5194/hess-15-3591-2011, 2011.

Smakhtin, V. Y. and Masse, B.: Continuous daily hydrograph simulation using duration curves of a precipitation index, Hydrological Processes, 14, 1083–1100, doi:10.1002/(SICI)1099-1085(20000430)14:6<1083::AID-HYP998>3.0.CO;2-2, 2000.

Smith, K., Barrett, C. B., and Box, P. W.: Participatory Risk Mapping for Targeting Research and Assistance: With an Example from East African Pastoralists, World Development, 28, 1945–1959, doi:10.1016/S0305-750X(00)00053-X, 2000.

Smith, M. B., Seo, D.-J., Koren, V. I., Reed, S. M., Zhang, Z., Duan, Q., Moreda, F., and Cong, S.: The distributed model intercomparison project (DMIP): motivation and experiment design, Journal of Hydrology, 298, 4–26, doi:10.1016/j.jhydrol.2004.03.040, 2004.

Smith, M. B., Koren, V., Reed, S., Zhang, Z., Zhang, Y., Moreda, F., Cui, Z., Mizukami, N., Anderson, E. A., and Cosgrove, B. A.: The distributed model intercomparison project - Phase 2: Motivation and design of the Oklahoma experiments, Journal of Hydrology, 418-419, 3–16, doi:10.1016/j.jhydrol.2011.08.055, 2012.

Stahl, K. and Demuth, S.: Linking streamflow drought to the occurrence of atmospheric circulation patterns, Hydrological Sciences Journal, 44, 467–482, doi:10.1080/02626669909492240, 1999.

Stahl, K., Hisdal, H., Hannaford, J., Tallaksen, L. M., van Lanen, H. A. J., Sauquet, E., Demuth, S., Fendekova, M., and Jódar, J.: Streamflow trends in Europe: evidence from a dataset of near-natural catchments, Hydrology and Earth System Sciences, 14, 2367–2382, doi:10.5194/hess-14-2367-2010, 2010.

Stahl, K., Tallaksen, L. M., Hannaford, J., and van Lanen, H. A. J.: Filling the white space on maps of European runoff trends: estimates from a multi-model ensemble, Hydrology and Earth System Sciences, 16, 2035–2047, doi:10.5194/hess-16-2035-2012, 2012.

Staudinger, M., Stahl, K., Seibert, J., Clark, M. P., and Tallaksen, L. M.: Comparison of hydrological model structures based on recession and low flow simulations, Hydrology and Earth System Sciences, 15, 3447–3459, doi:10.5194/hess-15-3447-2011, 2011.

Staudinger, M., Weiler, M., and Seibert, J.: Quantifying sensitivity to droughts - an experimental modeling approach, Hydrology and Earth System Sciences, 19, 1371–1384, doi:10.5194/hess-19-1371-2015, 2015.

Steinschneider, S. and Brown, C.: A semiparametric multivariate, multisite weather generator with low-frequency variability for use in climate risk assessments, Water Resources Research, 49, 7205–7220, doi:10.1002/wrcr.20528, 2013.

Stewart, I., Cayan, D., and Dettinger, M.: Changes toward earlier streamflow timing across western North America., Journal of Climate, pp. 1136–1155, doi:10.1175/JCLI3321.1, 2005.

Stoll, S., Hendricks Franssen, H. J., Butts, M., and Kinzelbach, W.: Analysis of the impact of climate change on groundwater related hydrological fluxes: a multi-model approach including different downscaling methods, Hydrology and Earth System Sciences, 15, 21–38, doi:10.5194/hess-15-21-2011, 2011.

Sunyer, M. A., Sørup, H. J. D., Christensen, O. B., Madsen, H., Rosbjerg, D., Mikkelsen, P. S., and Arnbjerg-Nielsen, K.: On the importance of observational data properties when assessing regional climate model performance of extreme precipitation, Hydrology and Earth System Sciences, 17, 4323–4337, doi:10.5194/hess-17-4323-2013, 2013.

Swiss Reinsurance: Weathering climate change: Insurance solutions for more resilient communities, Tech. rep., Swiss Re, Zurich, 2010.

Taylor, K. E., Stouffer, R. J., and Meehl, G. A.: An Overview of CMIP5 and the Experiment Design, Bulletin of the American Meteorological Society, 93, 485–498, doi:10.1175/BAMS-D-11-00094.1, 2012.

Tebaldi, C. and Arblaster, J. M.: Pattern scaling: Its strengths and limitations, and an update on the latest model simulations, Climatic Change, 122, 459–471, doi:10.1007/s10584-013-1032-9, 2014.

Tebaldi, C., Smith, R. L., Nychka, D., and Mearns, L. O.: Quantifying Uncertainty in Projections of Regional Climate Change: A Bayesian Approach to the Analysis of Multimodel Ensembles, Journal of Climate, 18, 1524–1540, doi:10.1175/JCLI3363.1, 2005.

Tebaldi, C., Arblaster, J. M., and Knutti, R.: Mapping model agreement on future climate projections, Geophysical Research Letters, 38, L23 701, doi:10.1029/2011GL049863, 2011.

Teng, J., Potter, N. J., Chiew, F. H. S., Zhang, L., Wang, B., Vaze, J., and Evans, J. P.: How does bias correction of regional climate model precipitation affect modelled runoff?, Hydrology and Earth System Sciences, 19, 711–728, doi:10.5194/hess-19-711-2015, 2015.

Teutschbein, C. and Seibert, J.: Bias correction of regional climate model simulations for hydrological climate-change impact studies: Review and evaluation of different methods, Journal of Hydrology, 456-457, 12–29, doi:10.1016/j.jhydrol.2012.05.052, 2012.

Teutschbein, C. and Seibert, J.: Is bias correction of regional climate model (RCM) simulations possible for non-stationary conditions?, Hydrology and Earth System Sciences, 17, 5061–5077, doi:10.5194/hess-17-5061-2013, 2013.

Themeßl, M. J., Gobiet, A., and Heinrich, G.: Empirical-statistical downscaling and error correction of regional climate models and its impact on the climate change signal, Climatic Change, pp. 449–468, doi:10.1007/s10584-011-0224-4, 2011a.

Themeßl, M. J., Gobiet, A., and Leuprecht, A.: Empirical-statistical downscaling and error correction of daily precipitation from regional climate models, International Journal of Climatology, 31, 1530–1544, doi:10.1002/joc.2168, 2011b.

Thrasher, B., Maurer, E. P., McKellar, C., and Duffy, P. B.: Technical Note: Bias correcting climate model simulated daily temperature extremes with quantile mapping, Hydrology and Earth System Sciences, 16, 3309–3314, doi:10.5194/hess-16-3309-2012, 2012.

Tobin, C., Nicotina, L., Parlange, M. B., Berne, A., and Rinaldo, A.: Improved Interpolation of Meteorological Forcings for Hydrologic Applications in a Swiss Alpine Region, Journal of Hydrology, pp. 77–89, doi:10.1016/j.jhydrol.2011.02.010, 2011.

Tschakert, P.: Views from the vulnerable: Understanding climatic and other stressors in the Sahel, Global Environmental Change, 17, 381–396, doi:10.1016/j.gloenvcha.2006.11.008, 2007.

Turco, M., Sanna, A., Herrera, S., Llasat, M. C., and Gutiérrez, J. M.: Large biases and inconsistent climate change signals in ENSEMBLES regional projections, Climatic Change, 120, 859–869, doi:10.1007/s10584-013-0844-y, 2013.

Uppala, S. M., K,Ñ´llberg, P. W., Simmons, A. J., Andrae, U., Bechtold, V. D. C., Fiorino, M., Gibson, J. K., Haseler, J., Hernandez, A., Kelly, G. A., Li, X., Onogi, K., Saarinen, S., Sokka, N., Allan, R. P., Andersson, E., Arpe, K., Balmaseda, M. A., Beljaars, A. C. M., Berg, L. V. D., Bidlot, J., Bormann, N., Caires, S., Chevallier, F., Dethof, A., Dragosavac, M., Fisher, M., Fuentes, M., Hagemann, S., Hólm, E., Hoskins, B. J., Isaksen, L., Janssen, P. A. E. M., Jenne, R., Mcnally, A. P., Mahfouf, J.-F., Morcrette, J.-J., Rayner, N. A., Saunders, R. W., Simon, P., Sterl, A., Trenberth, K. E., Untch, A., Vasiljevic, D., Viterbo, P., and Woollen, J.: The ERA-40 re-analysis, Quarterly Journal of the Royal Meteorological Society, 131, 2961–3012, doi:10.1256/qj.04.176, 2005.

van der Linden, P. and Mitchell, J., eds.: ENSEMBLES: Climate Change and its Impacts: Summary of research and results from the ENSEMBLES project, Met Office Hadley Centre, UK, 2009.

van der Schrier, G., van den Besselaar, E. J. M., Klein Tank, A. M. G., and Verver, G.: Monitoring European average temperature based on the E-OBS gridded data set, Journal of Geophysical Research: Atmospheres, 118, 5120–5135, doi:10.1002/jgrd.50444, 2013.

van Ulden, A., Lenderink, G., van den Hurk, B., and van Meijgaard, E.: Circulation statistics and climate change in Central Europe: PRUDENCE simulations and observations, Climatic Change, 81, 179–192, doi:10.1007/s10584-006-9212-5, 2007.

van Ulden, A. P. and van Oldenborgh, G. J.: Large-scale atmospheric circulation biases and changes in global climate model simulations and their importance for climate change in Central Europe, Atmospheric Chemistry and Physics, 6, 863–881, doi:10.5194/acp-6-863-2006, 2006.

Velázquez, J. A., Schmid, J., Ricard, S., Muerth, M. J., Gauvin St-Denis, B., Minville, M., Chaumont, D., Caya, D., Ludwig, R., and Turcotte, R.: An ensemble approach to assess hydrological models' contribution to uncertainties in the analysis of climate change impact on water resources, Hydrology and Earth System Sciences, 17, 565–578, doi:10.5194/hess-17-565-2013, 2013.

Vidale, P. L., Lüthi, D., Wegmann, R., and Schär, C.: European summer climate variability in a heterogeneous multi-model ensemble, Climatic Change, 81, 209–232, doi:10.1007/s10584-006-9218-z, 2007.

Viviroli, D., Mittelbach, H., Gurtz, J., and Weingartner, R.: Continuous simulation for flood estimation in ungauged mesoscale catchments of Switzerland - Part II: Parameter regionalisation and flood estimation results, Journal of Hydrology, 377, 208–225, doi:10.1016/j.jhydrol.2009.08.022, 2009a.

Viviroli, D., Zappa, M., Gurtz, J., and Weingartner, R.: An introduction to the hydrological modelling system PREVAH and its pre- and post-processing-tools, Environmental Modelling & Software, 24, 1209–1222, doi:10.1016/j.envsoft.2009.04.001, 2009b.

Vogel, C., Moser, S. C., Kasperson, R. E., and Dabelko, G. D.: Linking vulnerability, adaptation, and resilience science to practice: Pathways, players, and partnerships, Global Environmental Change, 17, 349–364, doi:10.1016/j.gloenvcha.2007.05.002, 2007.

Vogel, R. M. and Sankarasubramanian, A.: Validation of a watershed model without calibration, Water Resources Research, 39, doi:10.1029/2002WR001940, 2003.

von Storch, H. and Zwiers, F. W.: Statistical Analysis in Climate Research, Cambridge Univ. Press, UK, 2001.

Wagener, T.: Evaluation of catchment models, Hydrological Processes, 17, 3375–3378, doi:10.1002/hyp.5158, 2003.

Wang, C., Zhang, L., and Lee, S.: A global perspective on CMIP5 climate model biases, Nature Climate Change, 4, 201–205, doi:10.1038/NCLIMATE2118, 2014.

Weaver, C. P., Lempert, R. J., Brown, C., Hall, J. a., Revell, D., and Sarewitz, D.: Improving the contribution of climate model information to decision making: the value and demands of robust decision frameworks, Wiley Interdisciplinary Reviews: Climate Change, 4, 39–60, doi:10.1002/wcc.202, 2013.

Wetterhall, F., Pappenberger, F., He, Y., Freer, J., and Cloke, H. L.: Conditioning model output statistics of regional climate model precipitation on circulation patterns, Nonlinear Processes in Geophysics, 19, 623–633, doi:10.5194/npg-19-623-2012, 2012.

Whitfield, P. H.: Is Centre of Volume a robust indicator of changes in snowmelt timing?, Hydrological Processes, 27, 2691–2698, doi:10.1002/hyp.9817, 2013.

Wilby, R. L. and Dessai, S.: Robust adaptation to climate change, Weather, 65, 180–185, doi:10.1002/wea.543, 2010.

Wilby, R. L. and Harris, I.: A framework for assessing uncertainties in climate change impacts: Low-flow scenarios for the River Thames, UK, Water Resources Research, 42, W02 419, doi:10.1029/2005WR004065, 2006.

Wilby, R. L. and Quinn, N. W.: Reconstructing multi-decadal variations in fluvial flood risk using atmospheric circulation patterns, Journal of Hydrology, 487, 109–121, doi:10.1016/j.jhydrol.2013.02.038, 2013.

Wilcke, R. A. I., Mendlik, T., and Gobiet, A.: Multi-variable error correction of regional climate models, Climatic Change, 120, 871–887, doi:10.1007/s10584-013-0845-x, 2013.

Wisner, B., Blaikie, P., Cannon, T., and Davis, I.: At risk : natural hazards, people's vulnerability and disasters, Routledge, London, 2nd edn., 2004.

Wood, E. F., Roundy, J. K., Troy, T. J., van Beek, L. P. H., Bierkens, M. F. P., Blyth, E., de Roo, A., Döll, P., Ek, M., Famiglietti, J., Gochis, D., van de Giesen, N., Houser, P., Jaffé, P. R., Kollet, S., Lehner, B., Lettenmaier, D. P., Peters-Lidard, C., Sivapalan, M., Sheffield, J., Wade, A., and Whitehead, P.: Hyperresolution global land surface modeling: Meeting a grand challenge for monitoring Earth's terrestrial water, Water Resources Research, 47, W05 301, doi:10.1029/2010WR010090, 2011.

World Food Programme and Oxfam America: R4 Rural Resilience Initiative Quarterly Report, Tech. Rep. DeCeMBer, 2013.

Wu, C. F. J. and Hamada, M. S.: Experiments: Planning, Analysis, and Optimization, 2nd Edition, Wiley, URL http://www.wiley.com/WileyCDA/WileyTitle/productCd-0471699462.html, 2009.

Wynne, B.: May the sheep safely graze? A reflexive view of the expert-lay knowledge divide, in: Risk, Environment and Modernity: Towards a new ecology, edited by Lash, S., Szerszynski, B., and Wynne, B., pp. 44–83, Sage, London, 1996.

Zappa, M. and Gurtz, J.: Simulation of soil moisture and evapotranspiration in a soil profile during the 1999 MAP-Riviera Campaign, Hydrology and Earth System Sciences, 7, 903–919, doi:10.5194/hess-7-903-2003, 2003.

Zappa, M., Rotach, M. W., Arpagaus, M., Dorninger, M., Hegg, C., Montani, A., Ranzi, R., Ament, F., Germann, U., Grossi, G., Jaun, S., Rossa, A., and Vogt, S.: MAP D-PHASE : real-time demonstration of hydrological ensemble prediction systems, Atmospheric Science Letters, 9, 80–87, doi:10.1002/asl.183, 2008.

Zubler, E. M., Scherrer, S. C., Croci-Maspoli, M., Liniger, M. a., and Appenzeller, C.: Key climate indices in Switzerland; expected changes in a future climate, Climatic Change, 123, 255–271, doi:10.1007/s10584-013-1041-8, 2014.

Author contributions

This thesis consists of six studies presented in Chapters 2 to 7. The author contributions are as follows:

Study 1 Addor, N., Rössler, O., Köplin, N., Huss, M., Weingartner, R., and Seibert, J.: Robust changes and sources of uncertainty in the projected hydrological regimes of Swiss catchments, published in Water Resources Research, 2014.

I designed the study, coordinated the modeling experiment and analyzed the projections. O. Rössler, N. Köplin and myself produced the hydrological simulations (using WaSiM, PREVAH and HBV, respectively). M. Huss computed the glacier projections. J. Seibert helped with the organization of the modeling experiment. I was responsible for writing the paper, all authors contributed to the interpretation of the results and provided feedback on the manuscript.

Study 2 Addor, N., and Fischer, E. M.: The influence of natural variability and interpolation errors on bias characterization in RCM simulations, published in Journal of Geophysical Research: Atmospheres, 2015.

I performed the analyses, E. M. Fischer provided the multiple RCM runs to sample natural variability. I performed most of the analysis of the results and was responsible for writing the paper, with significant inputs from E. M. Fischer.

Study 3 Addor, N., Rohrer, M., Furrer, R., and Seibert, J.: Propagation of biases in climate models from the synoptic to the regional scale: implications for bias-correction, accepted pending revisions for publication in Journal of Geophysical Research: Atmospheres.

I retrieved and analyzed the RCM simulations, performed most of the analyses, except the circulation type classification, which was done by M. Rohrer. J. Seibert helped to develop the concept supporting the study. R. Furrer helped with the statistical analysis and interpretation. I was responsible for writing the paper, all authors contributed to the interpretation of the results and provided feedback on the manuscript.

Study 4 Addor, N., and Seibert, J.: Bias-correction for hydrological impact studies - beyond the daily perspective, published in Hydrological Processes, 2014.

I designed the study with J. Seibert. I performed the bias-correction, J. Seibert computed the correlation between the annual maximum discharge and the precipitation sum over the preceding days. We wrote the paper and analyzed the results together.

Study 5 Addor, N., Nikolova, S., and Seibert, J.: Trends in water balance as indicators of robustness for hydrological models in a changing climate, in preparation.

I designed the study and carried out the analyses. I supervised S. Nikolova during her Master's thesis, which inspired this study. I was responsible for writing the study and analyzing the results, with significant inputs from J. Seibert.

Study 6 Addor, N., Ewen, T., Johnson, L., Çöltekin, A., Derungs, C., and Muccione, V.: From products to processes: Academic events to foster interdisciplinary and iterative dialogue in a changing climate, published in Earth's Future, 2015.

I co-initiated the workshop-seminar series with T. Ewen and L. Johnson: we wrote together the proposal that was selected for funding through the Innovations Pool fund of the Department of Geography, University of Zurich. All authors contributed to the organisation of the workshop-seminar series and gave feedback on the manuscript. L. Johnson conducted the semantic analysis of participants' responses to the surveys. I was responsible for writing the paper, with important contributions from T. Ewen and L. Johnson.

Acknowledgements

My sincere gratitude goes to Jan Seibert. He accepted me as PhD student, although I had little experience in hydrological modeling under climate change, and virtually none in cross-country skiing. Over the four years of my PhD, he helped me to build a solid basis in hydrology, supported my participation in national projects and international meetings, introduced me to the scientific community, and allowed me to find my own research path. He also initiated me to cross-country skiing, and somehow convinced me to participate with him in a 50 km cross-country race during my first PhD year. He was then clearly faster. However, the following year, he was still taking his skis off when I crossed the finish line. So when the opportunity of sending me several thousands of miles away presented itself, he rapidly decided to let me finish my PhD in Oregon. Joking aside, these four years have been extremely enlightening and fun. Thank you Jan.

This thesis benefited from several collaborations, which were priceless opportunities to widen my perspective on climate change. I am particularly grateful to Tracy Ewen, Erich Fischer, Matthias Huss, Nina Köplin, Leigh Johnson, Marco Rohrer, Ole Rössler and Nadine Salzmann for their contribution and their encouragements, and for sharing their expertise with me.

Reinhard Furrer and Reto Knutti also played an important role in this thesis, I would like to thank them for their availability, for their support and for inspiring me through their research.

This thesis would not have been possible without the data, dedication and expertise of the Swiss Federal Office of Meteorology and Climatology, the Swiss Federal Office for the Environment, and the EU-funded projects ENSEMBLES and COST733.

I had the great opportunity to spend the second half of my PhD time at the University of Oregon, USA, which I would like to sincerely thank for hosting me.

I am grateful to the Swiss National Science Foundation for funding my PhD and supporting my stay in Oregon via a Doc.Mobility fellowship.

Before starting my PhD, I had had the chance to work with Thierry Lombardot, Simon Jaun, Massimiliano Zappa and Jacques Ambühl, who contributed to trigger my interest in scientific research, ensemble hydrological modeling and decision-making under uncertainty.

I truly enjoyed the time spent with my colleagues of the H2K group in Zurich and with the graduate students of the Department of Geological Sciences in Eugene.

Je remercie les membres de ma famille pour leur support infaillible et pour ces années passées ensemble.

I wonder what this PhD time would have been like without Georgie Bennett, who tirelessly kept faith in me, and was continuously convinced that one day, I would hand in a thesis I would be proud of. Once again you were right, Georgie. Thank you for your love and support.

Curriculum Vitae

Nans Addor

Research Applications Laboratory		Email:	naddor@ucar.edu
National Center for Atmospheric Research		Skype:	nansaddor
3450 Mitchell Lane		Phone:	(+1) 303 497 2873
Boulder, CO 80301, USA		Citizenship:	Swiss

Current and past positions

National Center for Atmospheric Research, Boulder, USA, 09.2015 - present.

Postdoctoral researcher within the Research Applications Laboratory, RAL, research on the relationship and trade-offs between complexity and fidelity in hydrological models.

University of Zurich, Switzerland, 06.2011 - 08.2015.

PhD student within the Hydrology and Climate Unit, H2K, research on uncertainties in the impacts of climate change on catchment hydrology and on bias-adjustment techniques for regional climate projections.

Education

PhD: University of Zurich, Switzerland, 06.2011 - 08.2015.

PhD thesis within the Hydrology and Climate Unit, H2K, Department of Geography: 'Impacts of climate change on discharge in Switzerland: Cascading uncertainties and robustness in models'.

Courtesy Research Assistant at the University of Oregon, USA, 08.2013 - 08.2015.

MSc: Swiss Federal Institute of Technology Zurich, ETHZ, 09.2007 - 12.2009.

ETHZ master's degree in Environmental Sciences, Major in Atmosphere and Climate.

Master's thesis at the Swiss Federal Institute for Forest, Snow and Landscape Research, WSL: 'Towards flood mitigation in the Sihl catchment using operational ensemble hydro-meteorological forecasts'.

BSc: Swiss Federal Institute of Technology Lausanne, EPFL, 10.2003 - 07.2006.

EPFL bachelor's degree in Environmental Sciences and Engineering.

Erasmus exchange year at the Polytechnic University of Madrid, 09.2005 - 07.2006.

Peer-reviewed journal articles

Addor, N. and Fischer, E. M.: The influence of natural variability and interpolation errors on bias characterization in RCM simulations, Journal of Geophysical Research: Atmospheres, 120, 180–10,195, doi:10.1002/2014JD022824, 2015

Addor, N., Ewen, T., Johnson, L., Çöltekin, A., Derungs, C., and Muccione, V.: From products to processes: Academic events to foster interdisciplinary and iterative dialogue in a changing climate, Earth's Future, 3, 289–297, doi:10.1002/2015EF000303, 2015

Addor, N., Rössler, O., Köplin, N., Huss, M., Weingartner, R., and Seibert, J.: Robust changes and sources of uncertainty in the projected hydrological regimes of Swiss catchments, Water Resources Research, 50, 7541–7562, doi:10.1002/2014WR015549, 2014

Addor, N. and Seibert, J.: Bias-correction for hydrological impact studies - beyond the daily perspective, Hydrological Processes, 28, 4823–4828, doi:10.1002/hyp.10238, 2014

Addor, N., Jaun, S., Fundel, F., and Zappa, M.: An operational hydrological ensemble prediction system for the city of Zurich (Switzerland): skill, case studies and scenarios, Hydrology and Earth System Sciences, 15, 2327–2347, doi:10.5194/hess-15-2327-2011, 2011

Lombardot, T., Kottmann, R., Giuliani, G., de Bono, A., Addor, N., and Glöckner, F. O.: MetaLook: a 3D visualisation software for marine ecological genomics, BMC Bioinformatics, 8, doi:10.1186/1471-2105-8-406, 2007

Articles in preparation or submitted

Addor, N., M. Rohrer, R. Furrer, and J. Seibert: Propagation of biases in climate models from the synoptic to the regional scale: implications for bias-correction, accepted pending revisions for publication in Journal of Geophysical Research: Atmospheres.

Addor, N., S. Nikolova, and J. Seibert: Trends in water balance as indicators of robustness for hydrological models in a changing climate, in preparation.

Academic activities and outreach

Co-lead author of the chapter 'Hydrological responses to climate change: river runoff and groundwater' of the CH2014-Impacts report 'Toward quantitative scenarios of climate change impacts in Switzerland'.

Co-initiator and co-organizer of the multi-day seminar 'Uncertainty in decision making in a changing climate' funded by the Department of Geography of University of Zurich and attended by 50 participants in March 2013.

Member of Kirsti Hakala's PhD committee, University of Zurich, thesis entitled 'A hydrological climate change impact assessment on Swiss catchments', 06.2015 - present.

Supervisor of Silviya Nikolova's master's thesis, University of Zurich, thesis entitled 'Robustness of hydrological simulations under contrasted climate conditions', 03.2013 - 01.2014.

Served as a reviewer for: Water Resources Research, Climatic Change, Hydrological Processes, Theoretical and Applied Climatology, Journal of Hydrology.

Assistant for the class ‘Hydrologische Modellierung’ (hydrological modeling), University of Zurich: preparation, explanation and correction of the exercises of four four-hour sessions, 09.2012 - 12.2012.

PhD courses attended

‘Watershed Science Master Class’ led by Jeff McDonnell, Biosphere 2, Arizona, USA, January 20-24, 2014.

‘Eawag Summer School in Environmental Systems Analysis’ led by Peter Reichert, Dübendorf, Switzerland, June 2-7, 2013.

‘The Water Cycle in a Changing Climate: Observations, Scenarios, Impacts’, 11th International NCCR Climate Summer School, Centro Stefano, Switzerland, September 9-14, 2012.

‘Uncertainty in Climate Change: An Integrated Approach, IMAGe Theme of the Year 2012’, led by Linda Mearns, National Center for Atmospheric Research, Boulder, USA, August 6-17, 2012.

‘Downscaling climate projections’, Technical University of Denmark, October 24-29, 2011.

‘Uncertainty in Environmental Modelling’ by Keith Beven, Uppsala University, Sweden, June 8-17, 2011.

Internships and civil service

NGO Nordesta: Civil service in the Brazilian rain forest and in Geneva to develop, promote and implement sustainable economic activities with and for communities of rubber tappers, 06.2010 - 05.2011.

Swiss Federal Institute for Forest, Snow and Landscape Research, WSL: Post-Master’s thesis research on flood ensemble prediction in the Sihl catchment, 03.2010 - 05.2010.

Swiss Federal Office of Meteorology and Climatology, MeteoSwiss: Comparative assessment of decision support schemes based on precipitation ensemble forecasts, 09.2008 - 03.2009.

Max Planck Institute for Marine Microbiology, MPI: Bioinformatic study of the relation between oceanic phosphate concentration and phosphate transporter genes in metagenomes, 05.2007 - 07.2007.

EPFL, Laboratory of Environmental Fluid Mechanics and Hydrology, EFLUM: Study on the influence of an urban environment on local meteorological parameters using a high-density station network distributed on the EPFL campus, 02.2007 - 04.2007.

United Nations Environment Programme, UNEP GRID-Europe: Processing of the NOAA-World Ocean Atlas 05 dataset to be displayed by an online Genes Mapserver, and development of interpolation and data mining tools, 09.2006 - 01.2007.

EPFL, Space Center: Study of potential remote sensing applications for a high altitude platform, 08.2006 - 09.2006.

International Union for Conservation of Nature, IUCN: Documentary research for an online article database promoting economic approaches for biodiversity conservation, 07.2006 - 08.2006.

Language skills

French: mother tongue, English: level C2, German: C1, Spanish: B2, Portuguese: B1.

Presentations and posters

Addor, N., Clark, M. P. and Nijssen, B.: Toward a systematic assessment of the relationship between the complexity and fidelity of hydrological models. Poster at the American Geophysical Union Fall Meeting (AGU), San Francisco, USA, December 16, 2015.

Addor, N., Rohrer, M., Furrer, R. and Seibert, J.: Exploring the links between biases in regional climate models and their representation of synoptic circulation types in the European Alps. Oral presentation at AGU, San Francisco, USA, December 19, 2014.

Ewen, T., Addor, N., Johnson, L., Coltekin, A., Derungs, C. and Muccione, V.: Fostering climate dialogue by introducing students to uncertainty in decision-making. Poster at AGU, San Francisco, USA, December 15, 2014.

Addor, N. and Fischer E. M.: Where does the mismatch between climate model simulations and observations come from? Differentiating between natural variability, interpolation errors and model biases. Poster at the 5th Annual Pacific Northwest Climate Science Conference, Seattle, USA, September 9, 2014.

Addor, N., Rössler O., Köplin N., Bernhard L., Bosshard T., Weingartner R. and Seibert J.: Bias-correction for climate impact studies, robust features and uncertainty sources in hydrological projections. Invited presentation at the Oregon Climate Change Research Institute, Corvallis, USA, May 7, 2014.

Ewen, T., Addor, N., Johnson, L., Coltekin, A., Derungs, C. and Muccione, V.: Creating dialogue: a workshop on 'Uncertainty in Decision Making in a Changing Climate'. Poster at the General Assembly of the European Geosciences Union (EGU), Vienna, Austria, April 2014.

Addor, N. and Seibert, J.: Robust changes and sources of uncertainty in the projected regimes of Swiss catchments. Oral presentation at the 3rd Annual UO Climate Change Research Symposium, University of Oregon, Eugene, USA, April 16, 2014.

Addor, N., Rössler O., Köplin N., Bernhard L., Bosshard T., Weingartner R. and Seibert J.: Robustness and uncertainties in future hydrological regimes of Swiss catchments. Poster at AGU, San Francisco, USA, December 2013.

Ewen, T., Addor, N., Johnson L., Coltekin, A., Derungs, C. and Muccione, V.: Creating dialogue: a workshop on 'Uncertainty in Decision Making in a Changing Climate', Poster at EGU, Vienna, April 2014.

Addor, N., Rössler O., Köplin N., Bernhard L., Bosshard T., Weingartner R. and Seibert J.: Robustness and and uncertainties in future hydrological regimes of Swiss catchments. Oral presentation at The Water Cycle in a Changing Climate Symposium, ETH Zurich, Switzerland, July 1, 2013.

Addor, N., Rössler O., Köplin N., Bernhard L., Bosshard T., Weingartner R. and Seibert J.: An uncertainty assessment of discharge projections for eight Swiss catchments. Oral presentation at EGU, Vienna, April 2013.

Addor, N. and Seibert J.: Can internal climate variability alone explain biases in regional climate simulations? Poster at EGU, Vienna, April 2013.

Tanaka, K. and Addor, N.: Permafrost-carbon feedbacks and climate stabilization costs. Poster at EGU, Vienna, April 2013.

Addor, N., Salzmann, N., Ewen, T. and Seibert, J.: Bias-correction conditioned on circulation patterns for ENSEMBLES precipitation simulations in mountainous regions. Oral presentation at EGU, Vienna, April 2012.

Addor, N., Jaun, S., Fundel, F. and Zappa, M.: An operational hydrological ensemble prediction system for the city of Zurich (Switzerland): assessing the added value of probabilistic forecasts. Poster at EGU, Vienna, April 2012.

Last updated: December 21, 2015